PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00 — A2

(11) International Publication Number: WO 98/55606

(43) International Publication Date: 10 December 1998 (10.12.98)

(21) International Application Number: PCT/CA98/00544

(22) International Filing Date: 2 June 1998 (02.06.98)

(30) Priority Data:

08/867,941	3 June 1997 (03.06.97)	US
09/074,658	8 May 1998 (08.05.98)	US

(71) Applicant *(for all designated States except US)*: CONNAUGHT LABORATORIES LIMITED [CA/CA]; 1755 Steeles Avenue West, North York, Ontario M2R 3T4 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LOOSMORE, Sheena, M. [CA/CA]; 70 Crawford Rose Drive, Aurora, Ontario L4G 4R4 (CA). DU, Run-Pan [CA/CA]; (CA). WANG, Quijun [CA/CA]; 299 Chelwood Drive, Thornhill, Ontario L4J 7Y8 (CA). YANG, Yan-Ping [CA/CA]; Apartment 709, 120 Torresdale Avenue, Willowdale, Ontario M2R 3N7 (CA). KLEIN, Michel, H. [CA/CA]; 16 Munro Boulevard, Willowdale, Ontario M2P 1B9 (CA).

(74) Agent: STEWART, Michael, I.; Sim & McBurney, 6th floor, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: LACTOFERRIN RECEPTOR GENES OF MORAXELLA

(57) Abstract

Purified and isolated nucleic acid molecules are provided which encode lactoferrin receptor proteins of *Moraxella*, such as *M. catarrhalis*, or a fragment or an analog of the lactoferrin receptor protein. The nucleic acid sequence may be used to produce recombinant lactoferrin receptor proteins Lbp1, Lbp2 or ORF3 of the strain of *Moraxella* free of other proteins of the *Moraxella* strain for purposes of diagnostics and medical treatment. Furthermore, the nucleic acid molecule may be used in the diagnosis of infection.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE OF INVENTION

LACTOFERRIN RECEPTOR GENES OF MORAXELLA

FIELD OF INVENTION

The present invention relates to the molecular cloning of genes encoding lactoferrin receptor (LfR) proteins and, in particular, to the cloning of lactoferrin binding protein genes (*lbp* genes) from *Moraxella (Branhamella) catarrhalis*.

REFERENCE TO RELATED APPLICATION

This application is a continuation-in-part of copending United States patent application No. 08/867,941 filed June 3, 1997.

BACKGROUND OF THE INVENTION

Moraxella (Branhamella) catarrhalis bacteria are Gram-negative diplococcal pathogens which are carried asymptomatically in the healthy human respiratory tract. However, in recent years, *M. catarrhalis* has been recognized as an important causative agent of otitis media. In addition, *M. catarrhalis* has been associated with sinusitis, conjunctivitis, and urogenital infections, as well as with a number of inflammatory diseases of the lower respiratory tract in children and adults, including pneumonia, chronic bronchitis, tracheitis, and emphysema (refs. 1 to 8). (Throughout this application, various references are cited in parentheses to describe more fully the state of the art to which this invention pertains. Full bibliographic information for each citation is found at the end of the specification, immediately preceding the claims. The disclosures of these references are hereby incorporated by reference into the present disclosure). Occasionally, *M. catarrhalis* invades to cause

septicaemia, arthritis, endocarditis, and meningitis (refs. 9 to 13).

M. catarrhalis colonizes the human upper respiratory tract and is an important cause of otitis media in infants and children as well as lower respiratory tract infections in adults with chronic obstructive pulmonary disease.

Otitis media is one of the most common illnesses of early childhood and approximately 80% of all children suffer at least one middle ear infection before the age of three (ref. 14). Chronic otitis media has been associated with auditory and speech impairment in children, and in some cases, has been associated with learning disabilities. Conventional treatments for otitis media include antibiotic administration and surgical procedures, including tonsillectomies, adenoidectomies, and tympanocentesis. In the United States, treatment costs for otitis media are estimated to be between one and two billion dollars per year.

In otitis media cases, *M. catarrhalis* is commonly co-isolated from middle ear fluid along with *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*, which are believed to be responsible for 50% and 30% of otitis media infections, respectively. *M. catarrhalis* is believed to be responsible for approximately 20% of otitis media infections (ref. 15). Epidemiological reports indicate that the number of cases of otitis media attributable to *M. catarrhalis* is increasing, along with the number of antibiotic-resistant isolates of *M. catarrhalis*. Thus, prior to 1970, no β-lactamase-producing *M. catarrhalis* isolates had been reported, but since the mid-seventies, an increasing number of β-lactamase-expressing isolates have been detected. Recent surveys suggest that up to

80 to 85% of clinical isolates produce β-lactamase (ref. 16, 22, 23).

Iron-restriction is a general host defence mechanism against microbial pathogens. A number of bacterial species including *Neisseria meningitidis* (ref. 17, 24), *N. gonorrhoeae* (ref. 25) and *M. catarrhalis* (ref. 17), express outer membrane proteins which specifically bind human lactoferrin.

M. catarrhalis infection may lead to serious disease. It would be advantageous to provide a recombinant source of lactoferrin binding proteins as antigens in immunogenic preparations including vaccines, carriers for other antigens and immunogens and the generation of diagnostic reagents. The genes encoding lactoferrin binding proteins and fragments thereof are particularly desirable and useful in the specific identification and diagnosis of *Moraxella* and for immunization against disease caused by *M. catarrhalis* and for the generation of diagnostic reagents.

SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules encoding a lactoferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the lactoferrin receptor protein. The nucleic acid molecules and isolated and purified lactoferrin binding proteins provided herein are useful for the specific detection of strains of *Moraxella* and for diagnosis of infection by *Moraxella*. The purified and isolated nucleic acid molecules provided herein, such as DNA, are also useful for expressing the *lbp* genes by recombinant DNA means for providing, in an economical manner, purified and isolated lactoferrin receptor

proteins free of other *Moraxella* proteins, as well as subunits, fragments or analogs thereof.

The lactoferrin receptor, subunits or fragments thereof or analogs thereof, as well as nucleic acid molecules encoding the same and vectors containing such nucleic acid molecules, are useful in immunogenic compositions for vaccinating against diseases caused by *Moraxella*, the diagnosis of infection by *Moraxella*, and as tools for the generation of immunological reagents.

Monoclonal antibodies or mono-specific antisera (antibodies) raised against the lactoferrin receptor protein produced in accordance with aspects of the present invention are useful for the diagnosis of infection by *Moraxella*, the specific detection of *Moraxella* (in, for example, *in vitro* and *in vivo* assays) and for the treatment of diseases caused by *Moraxella*.

In accordance with one aspect of the present invention, there is provided a purified and isolated nucleic acid molecule encoding a lactoferrin receptor protein of a strain of *Moraxella*, more particularly a strain of *M. catarrhalis*, specifically *M. catarrhalis* strain 4223, Q8 or VH19 or a fragment or an analog of the lactoferrin receptor protein. A fragment of the lactoferrin receptor protein is a portion of the protein which retains the immunological properties of the protein.

In one preferred embodiment of the invention, the nucleic acid molecule may encode only the Lbp1 protein of the *Moraxella* strain or only the Lbp2 protein of the *Moraxella* strain or only the ORF3 protein of the *Moraxella* strain. In another preferred embodiment of the invention, the nucleic acid may encode a fragment

of the lactoferrin receptor protein of a strain of *Moraxella* having a conserved amino acid sequence.

In a further aspect of the present invention, there is provided an isolated and purified nucleic acid molecule encoding at least one lactoferrin binding protein of *Moraxella* having a restriction map as shown in Figure 3 for *M. catarrhalis* 4223, Figure 5 for *M. catarrhalis* Q8 or Figure 17 for *M. catarrhalis* VH19 or the equivalent map from other strains of *Moraxella*.

In another aspect of the present invention, there is provided a purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of (a) a DNA sequence as set out in Figure 2 or 4 (SEQ ID Nos. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 69) or the complementary DNA sequence thereto; (b) a DNA sequence encoding an amino acid sequence as set out in Figure 2 or 4 (SEQ ID Nos. 11, 12, 13, 14, 15, 16, 17, 18, 70) or the complementary DNA sequence thereto; and (c) a DNA sequence encoding a functional lactoferrin receptor preotein of *Moraxella*, which may be a DNA sequence which hybridizes under stringent conditions to any one of the DNA sequences defined in (a) or (b). The DNA sequence defined in (c) may have at least about 90% sequence identity with any one of the DNA sequences defined in (a) or (b). Stringent conditions of hybridization are described below. Sequence identity is determined in the manner described below.

In an additional aspect, the present invention includes a vector adapted for transformation of a host, comprising a nucleic acid molecule as provided herein and may have the characteristics of a nucleotide sequence contained within vectors pLD3, pLDW3, pLD1-8 and pLDW1.

The vector may be adapted for expression of the encoded lactoferrin receptor protein, fragments or

analogs thereof, in a heterologous or homologous host, in either a lipidated or non-lipidated form. Accordingly, a further aspect of the present invention provides an expression vector adapted for transformation of a host comprising a nucleic acid molecule as provided herein and expression means operatively coupled to the nucleic acid molecule for expression by the host of the lactoferrin receptor protein or the fragment or analog of the lactoferrin receptor protein.

In specific embodiments of this aspect of the invention, the nucleic acid molecule may encode substantially all the lactoferrin receptor protein, only the Lbp1 protein of the *Moraxella* strain, only the Lbp2 protein of the *Moraxella* strain, only the ORF3 protein of the *Moraxella* strain, or fragments of the Lbp1, Lbp2 or ORF3 proteins.

The expression means may include a nucleic acid portion encoding a leader sequence for secretion from the host of the lactoferrin receptor protein or the fragment or the analog of the lactoferrin receptor protein. The expression means also may include a nucleic acid portion encoding a lipidation signal for expression from the host of a lipidated form of the lactoferrin receptor protein or the fragment or the analog of the lactoferrin receptor protein. The host may be selected from, for example, *Escherichia coli, Bacillus, Bordetella, Haemophilus, Moraxella,* fungi, yeast or baculovirus and Semliki Forest virus expression system may be used. In a particular embodiment, the plasmid adapted for expression of Lbp2 is pRD2A, pRD2B, pQW2A or pQW2B; the plasmid adapted for expression of Lbp1 is pRD1A, pRD1B, PQ1A or pQ1B; and the plasmid adapted for expression of ORF3 is pLRD3 or pLQW3.

In an additional aspect of the invention, there is provided a transformed host containing an expression vector as provided herein. The invention further includes a recombinant lactoferrin receptor protein or fragment or analog thereof of a strain of *Moraxella* producible by the transformed host.

Such recombinant lactoferrin receptor protein may be provided in substantially pure form according to a further aspect of the invention, which provides a method of forming a substantially pure recombinant lactoferrin receptor protein, which comprises growing the transformed host provided herein and isolating and purifying the lactoferrin receptor protein, analog or fragment thereof. The lactoferrin receptor protein may be expressed in inclusion bodies, which may be purified free from cellular material and soluble proteins and lactoferrin receptor protein solubilized from the purified inclusion bodies, and the lactoferrin receptor protein purified free from other solubilized materials. The substantially pure recombinant lactoferrin receptor protein may comprise Lbp1 alone, Lbp2 alone, ORF3 or a mixture of two or more of such proteins. The recombinant protein is generally at least about 70% pure, preferably at least about 90% pure.

Further aspects of the present invention, therefore, provide recombinantly-produced Lbp1 protein (or a fragment or analog thereof) of a strain of *Moraxella* devoid of the Lbp2 and ORF3 proteins of the *Moraxella* strain and any other protein of the *Moraxella* strain, recombinantly-produced Lbp2 protein (or a fragment or analog thereof) of a strain of *Moraxella* devoid of the Lbp1 and ORF3 proteins of the *Moraxella* strain and any other protein of the *Moraxella* strain, and recombinantly-produced ORF3 protein (or a fragment or analog thereof) of a strain of *Moraxella* devoid of

the Lbp1 and Lbp2 proteins of the *Moraxella* strain and any other protein of the *Moraxella* strain. The *Moraxella* strain may be *M. catarrhalis* 4223, Q8 or VH19 strain.

The invention further includes, in an additional aspect, an open reading frame protein 3 (ORF3) of a *Moraxella* strain or a fragment or analog of the lactoferrin binding protein which is encoded by region downstream from the genes encoding Lbp2 and Lbp1 proteins of the *Moraxella* strain. The ORF3 may be from a strain of *M. catarrhalis*, which may be strain 4223 or Q8. The Lbp3 may have a molecular mass of about 60 kDa.

In accordance with another aspect of the invention, an immunogenic composition is provided which comprises at least one active component selected from at least one nucleic acid molecule as provided herein, at least one recombinant protein as provided herein or at least one novel protein as provided herein, and a pharmaceutically acceptable carrier therefor or vector therefor. The at least one active component produces an immune response when administered to a host.

The immunogenic compositions provided herein may be formulated as a vaccine for *in vivo* administration to a host to provide protection against disease caused by *M. catarrhalis*. For such purpose, the compositions may be formulated as a microparticle, capsule, ISCOM or liposome preparation. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. The immunogenic compositions of the invention (including vaccines) may further comprise at least one other immunogenic or immunostimulating material and the immunostimulating material may be at least one adjuvant or at least one cytokine.

Suitable adjuvants for use in the present invention include (but are not limited to) aluminum phosphate, aluminum hydroxide, QS21, Quil A, derivatives and components thereof, ISCOM matrix, calcium phosphate, calcium hydroxide, zinc hydroxide, a glycolipid analog, an octadecyl ester of an amino acid, a muramyl dipeptide, polyphosphazene, ISCOPREP, DC-chol, DDBA and a lipoprotein and other adjuvants to induce a TH1 response. Advantageous combination of adjuvants are described in copending United States Patent Applications No. 08/261,194 filed June 16, 1994 and 08/483,856 filed June 7, 1995, assigned to the assignee hereof and the disclosure of which is incorporated herein by reference (WO 95/34308, published November 21, 1995).

In accordance with another aspect of the invention, there is provided a method for generating an immune response in a host, comprising the step of administering to a susceptible host, such as a human, an effective amount of the immunogenic composition as recited above. The immune response may be humoral or a cell-mediated immune response and may provide protection against disease caused by *Moraxella*. Hosts in which protection against disease may be conferred include primates, including humans.

In a further aspect, there is provided a live vector for delivery of lactoferrin receptor to a host, comprising a vector containing the nucleic acid molecule as described above. The vector may be selected from *Salmonella*, *Mycobacterium bovis*, BCG, adenovirus, poxvirus, vaccinia and poliovirus.

The nucleic acid molecules provided herein are useful in diagnostic applications. Accordingly, in a further aspect of the invention, there is provided a method of determining the presence, in a sample, of

nucleic acid encoding a lactoferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

a) contacting the sample with a nucleic acid molecule as provided herein to produce duplexes comprising the nucleic acid molecule and any nucleic acid molecule encoding the lactoferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

b) determining the production of the duplexes.

In addition, the present invention provides a diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a lactoferrin receptor protein of a strain of *Moraxella*, comprising:

a) a nucleic acid molecule as provided herein;

b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any such nucleic acid present in the sample and hybridizable with the nucleic acid present in the sample and hydridizable with the nucleic acid molecule; and

c) means for determining production of the duplexes.

The invention further includes the use of the nucleic acid molecules and proteins provided herein as medicines. The invention additionally includes the use of the nucleic acid molecules and proteins provided herein in the manufacture of medicaments for protection against disease caused by strains of *Moraxella*.

Advantages of the present invention include:

- an isolated and purified nucleic acid molecule encoding a lactoferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the lactoferrin receptor protein;
- recombinantly-produced lactoferrin receptor proteins, including Lbp1, Lbp2 and ORF3 and fragments

and analogs thereof free from each other and other *Moraxella* proteins;

- open reading frame protein 3; and
- diagnostic kits and immunological reagents for specific identification of *Moraxella*.

BRIEF DESCRIPTION OF DRAWINGS

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1 shows partial sequence of the 2.2 kb PCR amplified fragments of the *lbpA* genes from *M. catarrhalis* 4223 or Q8, which were used to probe the phage libraries. In the figure, Tbp1 is the deduced 4223 Tbp1 sequence (as described in United States Patent Application No. 08/613,009 filed March 8, 1996, assigned to the assignee hereof and the disclosure of which is incorporated herein by reference) (SEQ ID No: 19), Lbp1 is the deduced full-length 4223 Lbp1 sequence (SEQ ID No: 3) used here solely for aligning the PCR fragments, PCR4 is the 4223 PCR fragment (SEQ ID No: 20), and PCR5 is a partial sequence of the Q8 PCR fragment (SEQ ID No: 21). Only single strand sequence was obtained for the PCR fragments and "X" has been inserted where there was a doubtful sequence. Dashes have been used for maximum alignment. The underlined sequence in Lbp1 (MVQYTYRKGKENKAH - SEQ ID No: 22) represents the position of a CNBr peptide used to generate the 5'-PCR primer.

Figure 2 shows the nucleotide (SEQ ID No: 1, full sequence; SEQ ID No: 2, Lbp2 coding sequence; SEQ ID No: 3, Lbp1 coding sequence, first methionine; SEQ ID No: 4, Lbp1 coding sequence, second methionine; SEQ ID No: 5, ORF3 coding sequence) and deduced amino acid sequences (SEQ ID No: 11, Lbp2; SEQ ID No: 12, Lbp1, first methionine; SEQ ID No: 13, Lbp1, second

methionine; SEQ ID No: 14, ORF3) of the putative *lfr* locus from *M. catarrhalis* 4223. There are three tandem genes in the putative *lfr* locus identified as *lbpB*, *lbpA* and *orf3*. Potential promoter elements found upstream of the *lbpB* and *lbpA* genes are indicated by underlining.

Figure 3 shows a restriction map of clone pLD1-8 containing the *lbpA*, *lbpB*, and *orf3* genes from *M. catarrhalis* isolate 4223.

Figure 4 shows the nucleotide (SEQ ID No: 6, full sequence; SEQ ID No: 7, Lbp2 coding sequence; SEQ ID No: 8, Lbp1 coding sequence, first methionine; SEQ ID No: 9, Lbp2, second methionine; SEQ ID No: 10, ORF3 coding sequence) and deduced amino acid sequences (SEQ ID No: 15, Lbp2; SEQ ID No: 16, Lbp1, first methionine; SEQ ID No: 17, Lbp1, second methionine; SEQ ID No: 18, Lbp3) of the putative *lfr* locus from *M. catarrhalis* Q8. There are three tandem genes in the putative *lfr* locus identified as *lbpB*, *lbpA* and *orf3*. Potential promoter elements found upstream of the *lbpB* and *lbpA* genes are indicated by underlining.

Figure 5 shows a restriction map of clone pLDW1 containing the *lbpA*, *lbpB* and *orf3* genes from *M. catarrhalis* isolate Q8.

Figure 6 shows a comparison of the amino acid sequences of Lbp1 from *M. catarrhalis* strains 4223 (SEQ ID No: 12) and Q8 (SEQ ID No: 16), *N. meningitidis* strains BNCV (SEQ ID No: 23) and H44/76 (SEQ ID No: 75), and *N. gonorrhoeae* strain FA19 (SEQ ID No: 24). Also shown is the partial carboxy terminal sequence of Lbp2 from *N. meningitidis* strains BNCV (SEQ ID No: 76) and H44/76 (SEQ ID No: 77) and *N. gonorrhoeae* strain FA19 (SEQ ID No: 88). Dots indicate identical residues

and dashes have been introduced to achieve maximum sequence alignment.

Figure 7 shows a comparison of the amino acid sequences of Lbp2 from *M. catarrhalis* strains 4223 (SEQ ID No: 11), Q8 (SEQ ID No: 15) and VH19 (SEQ ID No: 70). Dots indicate identical residues. The arrow indicates the lipidated cysteine of a potential mature Lbp2 lipoprotein. The residues conserved with Tbp2 proteins are underlined and the RGD sequence is italicized.

Figure 8 shows a comparison of the amino acid sequences of Tbp2 (USPA No: 08/613,009) (SEQ ID No: 25) and Lbp2 from *M. catarrhalis* strain 4223 (SEQ ID No: 11). Dots indicate identical residues and dashes have been inserted to achieve maximum sequence alignment. The asterisks indicate conserved residues and the putative site of lipidation for both proteins is indicated by the arrow.

Figure 9 shows a comparison of the amino acid sequences of ORF3 from *M. catarrhalis* strains 4223 (SEQ ID No: 14) and Q8 (SEQ ID No: 18). Dots indicate identical residues and dashes have been introduced for maximum alignment.

Figure 10 shows the construction of plasmids for expression of recombinant Lbp1 protein from *E. coli*. Plasmids pRD1A and pRD1B express 4223 Lbp1 from the first or second methionine residues, respectively. Plasmids pQW1A and pQW1B express Q8 Lbp1 from the first or second methionine residues, respectively.

Figure 11, comprising panels A and B, shows the expression of recombinant Lbp1 (rLbp1) proteins from *E. coli*. Panel A shows the expression of the Q8 Lbp1 proteins and panel B shows the expression of the 4223 Lbp1 proteins. Lane 1, molecular weight marker. Lanes 2 and 3 demonstrate the induced expression of the

longer Lbp1 starting from the first methionine residues and lanes 4 and 5 illustrate the expression of the shorter Lbp1 proteins starting from the second methionine residues. Lanes 6, 7, 8 and 9 are uninduced samples.

Figure 12 shows the construction of plasmids for expression of recombinant Lbp2 (rLbp2) protein from *E. coli*. Plasmids pRD2A and pRD2B express 4223 Lbp2 with or without the native leader sequence, respectively. Plasmids pQW2A and pQW2B express Q8 Lbp2 with or without the native leader sequence, respectively.

Figure 13 shows the construction of a plasmid for expression of recombinant ORF3 (rORF3) proteins from *E. coli*.

Figure 14 shows a purification scheme for rLbp1 expressed from *E. coli*.

Figure 15 shows an SDS PAGE gel of the purification of Q8 Lbp1 from *E. coli*. Lane 1, BL21(DE3) lysate; lane 2, soluble proteins after 50 mM Tris/5 mM AEBSF/0.5 M NaCl, pH 8.0 extraction; lane 3, soluble proteins after 50 mM Tris/0.5% Triton X-100/10 mM EDTA, pH 8.0 extraction; lane 4, soluble proteins after 50 mM Tris-HCl/1% octylglucoside, pH 8.0 extraction; lane 5, solubilized inclusion bodies; lane 6, purified Lbp1.

Figure 16 shows the nucleotide sequence (SEQ ID No: 69) of the *M. catarrhalis* strain VH19 *lbpB* gene and the deduced amino acid sequence (SEQ ID No: 70) of the corresponding Lbp2 protein.

Figure 17 shows a partial restriction map of the *M. catarrhalis* strain VH19 *lbpB* gene.

Figure 18, comprising panels A, B and C, shows SDS-PAGE gels of the purification of recombinant Lbp proteins. Panel A shows an SDS-PAGE gel of the purification of Q8 rLbp1. Panels B and C show the

purification of Q8 rLbp2 and 4223 rLbp2, respectively. Lane 1, molecular weight markers; lane 2, whole cell lysates; lane 3, inclusion bodies; lane 4, purified protein.

Figure 19, comprising panels A and B, shows binding of recombinant Lbp proteins to lactoferrin. Panel A shows an SDS PAGE gel of purified recombinant proteins. Panel B shows the binding of recombinant proteins to human lactoferrin. Lane 1, molecular weight markers; lane 2, Q8 rLbp1; lane 3, Q8 rLbp2; lane 4, 4223 rLbp2.

Figure 20, comprising panels A, B and C, shows an immunoblot of *M. catarrhalis* strains reacted with anti-rLbp1 and anti-rLbp2 antibodies. Panel A: whole cell lysates probed with anti-Q8 rLbp1 + anti-Q8 rLbp2 antisera. All cells were grown in the presence of EDDA. Panel B: whole cell lystaes probed with anti-Q8 rLbp1 antibody. Panel C: whole cell lysates probed with anti-Q8 rLbp2 antibody. Lane 1, strain Q8; lane 2, strain 4223; lane 3, strain VH19; lane 4, strain LES-1; lane 5, strain H-04; lane 6, strain 3. + indicates growth in the presence of EDDA and - indicates growth in the absence of EDDA.

GENERAL DESCRIPTION OF THE INVENTION

Any *Moraxella* strain may be conveniently used to provide the purified and isolated nucleic acid, which may be in the form of DNA molecules, comprising at least a portion of the nucleic acid coding for a lactoferrin receptor as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection.

In this application, the terms "lactoferrin receptor" (LfR) and "lactoferrin binding proteins" (Lbp) are used to define a family of Lbp1, Lbp2 and/or ORF3 proteins which includes those having variations in their amino acid sequences including those naturally occurring in various strains of, for example, *Moraxella*. The purified and isolated DNA molecules comprising at least a portion coding for lactoferrin receptor of the present invention also includes those encoding functional analogs of lactoferrin receptor proteins Lbp1, Lbp2 and/or Lbp3 of *Moraxella*. In this application, a first protein is an "analog" of a second protein if the first protein is immunologically related to and/or has the same function as the second protein. The analog may be, for example, a substitution, addition or deletion mutant thereof.

Lactoferrin receptor proteins were purified from *M. catarrhalis* membrane preparations by affinity chromatography on biotinylated human lactoferrin. Cyanogen bromide fragments were generated and amino acid sequence analysis of a 13 kDa fragment provided an internal Lbp1 sequence of MVQYTYRKGKENKAH (SEQ ID No: 22) underlined in Figure 6. The C-terminus of *M. catarrhalis* Tbp1 (United States Patent Application No. 08/613,009), *N. meningitidis* Tbp1 (ref. 27) and *H. influenzae* Tbp1 (ref. 31) has a conserved LEMKF (SEQ ID No: 26) sequence. Oligonucleotide primers were generated based upon these two sequences and used to PCR amplify an approximately 2.2kb fragment of the *lbpA* gene from *M. catarrhalis* strains 4223, Q8 and VH19. Partial sequence analysis demonstrated that the amplified genes were *lbpA* and not *tbpA* (see Fig. 1). The 2.2 kb PCR fragments were used to screen genomic libraries.

Chromosomal DNA from 4223, Q8 and VH19 was partially digested with Sau3A I and 15 to 23 kb fragments were purified before cloning into BamH I arms of the lambda vector EMBL3. The libraries were screened with the PCR fragment and positive clones were subjected to three rounds of plaque purification. Phage clone 4223LfR.17 containing an approximately 16 kb insert from 4223 and phage clone Q8LfR.13 containing an approximately 16 kb insert from Q8 were selected for further analysis.

Restriction enzyme and Southern blot analyses revealed that an internal Hind III fragment of approximately 9 kb contained at least a portion of the *lbpA* gene for both phage clones. The approximately 9 kb Hind III fragments were subcloned into pUC or pBluescript-based plasmids and sequenced. In each case, they contained the complete *lbpA* gene as well as an upstream gene identified as *lbpB*, and a downstream gene designated as *orf3*. The *lbpB-lbpA* gene arrangement is the same as present for *Neisseria* strains, but there has been no identification of a third gene for these organisms.

The gene arrangement is different than that observed for the *M. catarrhalis tfr* operon which was *tbpA-orf-tbpB* (United States Patent Application No. 08/613,009). There are promoter elements found upstream of both the *lbpB* and *lbpA* genes from strains 4223 and Q8. The third ORF is located immediately downstream of *lbpA*, separated by a single nucleotide.

By analogy with the *N. meningitidis* and *N. gonorrhoeae* transferrin receptor operons (ref. 26, 27, 28), the lactoferrin receptor operon was presumed to consist of two genes encoding lactoferrin binding proteins 1 and 2 (Lbp1 and Lbp2) (ref. 29). However,

we report here that, for *M. catarrhalis*, there also appears to be a third gene located immediately downstream of *lbpA* encoding a potential lactoferrin binding protein 3 (ORF3).

The *M. catarrhalis* 4223 and Q8 *lbpA* genes encode proteins of molecular mass about 110 kDa and that are highly conserved with only seven residues difference between them. The N-terminal sequence of the native Lbp protein is unknown and there are two possible ATG start codons at positions 1 or 16. The first of these is adjacent to consensus sequences for promoter elements and the second is followed by a putative signal sequence. The exact peptide sequence used to design the PCR amplification primers was not found. When compared with other known Lbp1 sequences from *N. meningitidis* (refs. 31, 24) or *N. gonorrhoeae* (ref. 25) there is about 32% sequence identity and about 50% sequence homology between the *M. catarrhalis* and the *Neisseria* proteins. There is some homology between the *M. catarrhalis* Lbp1 and Tbp1 proteins as shown in Figure 1, but it is very scattered.

The *M. catarrhalis* 4223, Q8 and VH19 *lbpB* genes encode 898, 894 and 906 amino acid proteins, respectively. The *M. catarrhalis* Lbp2 proteins from strains 4223 and Q8 are 92% identical and 95% homologous while that from VH19 is 77% identical and 84% similar to the 4223 and Q8 Lbp2 proteins (Figure 7). There is a consensus sequence for lipidation at the Cys^{32} residue, suggesting that Lbp2 is a lipoprotein like Tbp2. There is little homology between the *M. catarrhalis* Lbp2 and Tbp2 proteins (Fig. 8) with the exception of a previously identified peptide sequence (LEGGFY (SEQ ID No: 27)) that is also found in *N. meningitidis* and *H. influenzae* Tbp2 (ref. 30).

The sequence of the proposed *M. catarrhalis lfr*-related downstream *orf3* is conserved between strains 4223 and Q8. The encoded 4223 and Q8 ORF3 proteins when compared to the PIR and Swiss Prot protein databases were found to be previously unknown. The ORF3 protein may bind lactoferrin itself or may be an associated or regulatory protein for Lbp1 and/or Lbp2.

Expression vectors have been assembled from the *lbpA* and *lbpB* genes and recombinant Lbp1 and Lbp2 proteins isolated and purified, as described in detail in the Examples below.

Results shown in Table 1 below illustrate the ability of anti-Lbp1 guinea pig antiserum, produced by immunization with affinity purified Lbp1, to lyse *M. catarrhalis*. The results show that the antisera produced by immunization with Lbp1 protein isolated from *M. catarrhalis* isolate 4223 was bactericidal against a homologous non-clumping *M. catarrhalis* strain RH408 (a strain previously deposited in connection with United States Patent Application No. 08/328,589, assigned to the assignee hereof (WO 96/12733 published May 2, 1996)) derived from isolate 4223. In addition, antisera produced by immunization with Lbp1 protein isolated from *M. catarrhalis* 4223 were bactericidal against the heterologous non-clumping strain Q8. The results in Table 3 show that similarly-produced anti-Lbp2 guinea pig antiserum was bactericidal for the homologous strain and for three of five hetrologous strains. The ability of isolated and purified lactoferrin binding protein to generate bactericidal antibodies is *in vivo* evidence of utility of these proteins as vaccines to protect against disease caused by *Moraxella*.

Thus, in accordance with another aspect of the present invention, there is provided a vaccine against *Moraxella* comprising an immunogenically-effective amount of lactoferrin binding protein or fragment or analog thereof, or a nucleic acid molecule (DNA or RNA) encoding the lactoferrin binding protein or fragment or analog thereof, and a physiologically-acceptable carrier therefor. The lactoferrin binding protein or fragment or analog thereof provided herein may also be used as a carrier protein for haptens, polysaccharide or peptides to make conjugate vaccines against antigenic determinants unrelated to lactoferrin binding proteins.

In additional embodiments of the present invention, therefore, the lactoferrin binding protein as provided herein may be used as a carrier molecule to prepare chimeric molecules and conjugate vaccines (including glycoconjugates) against pathogenic bacteria, including encapsulated bacteria. Thus, for example, glycoconjugates of the present invention may be used to confer protection against disease and infection caused by any bacteria having polysaccharide antigens including lipooligosaccharides (LOS) and PRP. Such bacterial pathogens may include, for example, *Haemophilus influenzae, Streptococcus pneumoniae, Escherichia coli, Neisseria meningitidis, Salmonella typhi, Streptococcus mutans, Cryptococcus neoformans, Klebsiella, Staphylococcus aureus* and *Pseudomonas aeruginosa*. Particular antigens which can be conjugated to lactoferrin binding protein and methods to achieve such conjugations are described in U.S. Patent Application No. 08/433,522 filed November 23, 1993 (WO 94/12641), assigned to the assignee hereof and the disclosure of which is hereby incorporated by reference thereto.

In another embodiment, the carrier function of lactoferrin binding protein may be used, for example, to induce an immune response against abnormal polysaccharides of tumour cells, or to produce anti-tumour antibodies that can be conjugated to chemotherapeutic or bioactive agents.

The lactoferrin binding protein provided herein is useful as a diagnostic reagent, as an antigen or for the generation of anti-lactoferrin protein binding antibodies, antigen for vaccination against disease caused by species of *Moraxella* and for detecting infection by *Moraxella* and other such bacteria.

The invention extends to lactoferrin binding proteins or fragments or analogs thereof or nucleic acid molecules encoding the same from *Moraxella catarrhalis* for use as an active ingredient in a vaccine against disease caused by infection with *Moraxella*. The invention also extends to a pharmaceutical vaccinal composition containing lactoferrin binding proteins or fragments or analogs thereof or nucleic acid molesules encoding the same from *Moraxella catarrhalis* and optionally, a pharmaceutically acceptable carrier and/or diluent.

In a further aspect the invention provides the use of lactoferrin binding proteins or fragments or analogs thereof or nucleic acid molesules encoding the same for the preparation of a pharmaceutical vaccinal composition for immunization against disease caused by infection with *Moraxella*.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, treatment of, for example, *Moraxella* infections and the generation of immunological and other

diagnostic reagents. A further non-limiting discussion of such uses is further presented below.

1. Vaccine Preparation and Use

Immunogenic compositions, suitable to be used as vaccines, may be prepared from immunogenic lactoferrin receptor proteins, analogs and fragments thereof encoded by the nucleic acid molecules as well as the nucleic acid molecules disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-lactoferrin receptor antibodies and antibodies that are opsonizing or bactericidal. Should the vaccinated subject be challenged by *Moraxella*, the antibodies bind to the lactoferrin receptor and thereby prevent access of the bacteria to an iron source which is required for viability. Furthermore, opsonizing or bactericidal anti-lactoferrin receptor antibodies may also provide protection by alternative mechanisms.

Immunogenic compositions, including vaccines, may be prepared as injectables, as liquid solutions or emulsions. The lactoferrin receptor proteins, analogs and fragments thereof and encoding nucleic acid molecules as well as the nucleic acid molecules described herein may be mixed with pharmaceutically acceptable excipients which are compatible with the lactoferrin receptor proteins, fragments, analogs or nucleic acid molecules. Such excipients may include water, saline, dextrose, glycerol, ethanol, and combinations thereof. The immunogenic compositions and vaccines may further contain auxiliary substances, such as wetting or emulsifying agents, pH buffering agents, or adjuvants, to enhance the effectiveness of the vaccines. Immunogenic compositions and vaccines may be administered parenterally, by injection subcutaneously, intradermally or intramuscularly. Alternatively, the immunogenic compositions provided according to the

present invention, may be formulated and delivered in a manner to evoke an immune response at mucosal surfaces. Thus, the immunogenic composition may be administered to mucosal surfaces by, for example, the nasal or oral (intragastric) routes. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. Some such targeting molecules include vitamin B12 and fragments of bacterial toxins, as described in WO 92/17167 (Biotech Australia Pty. Ltd.), and monoclonal antibodies, as described in U.S. Patent No. 5,194,254 (Barber et al). Alternatively, other modes of administration, including suppositories and oral formulations, may be desirable. For suppositories, binders and carriers may include, for example, polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example, pharmaceutical grades of saccharine, cellulose and magnesium carbonate. These compositions may take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 1 to 95% of the lactoferrin receptor proteins, fragments, analogs and/or nucleic acid molecules.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective, protective and immunogenic. The quantity to be administered depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and, if needed, to produce a cell-mediated immune response. Precise amounts of active ingredient required to be administered depend on the judgement of the practitioner. However, suitable dosage ranges are readily determinable by one skilled in

the art and may be of the order of micrograms of the lactoferrin receptor proteins, analogs and fragments thereof and/or nucleic acid molecules. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and will vary according to the size of the host.

The nucleic acid molecules encoding the lactoferrin receptor of *Moraxella* may be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus containing the nucleic acid molecules. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system is contained in, for example, O'Hagan (ref. 18). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulmer et al. (ref. 19).

Immunogenicity can be significantly improved if the antigens are co-administered with adjuvants, commonly used as an 0.05 to 1.0 percent solution in phosphate - buffered saline. Adjuvants enhance the immunogenicity of an antigen but are not necessarily immunogenic themselves. Adjuvants may act by retaining the antigen locally near the site of administration to produce a depot effect facilitating a slow, sustained release of antigen to cells of the immune system. Adjuvants can also attract cells of the immune system to an antigen depot and stimulate such cells to elicit immune responses.

Immunostimulatory agents or adjuvants have been used for many years to improve the host immune responses to, for example, vaccines. Intrinsic adjuvants, such as lipopolysaccharides, normally are the components of killed or attenuated bacteria used as vaccines. Extrinsic adjuvants are immunomodulators which are typically non-covalently linked to antigens and are formulated to enhance the host immune responses. Thus, adjuvants have been identified that enhance the immune response to antigens delivered parenterally. Some of these adjuvants are toxic, however, and can cause undesirable side-effects, making them unsuitable for use in humans and many animals. Indeed, only aluminum hydroxide and aluminum phosphate (collectively commonly referred to as alum) are routinely used as adjuvants in human and veterinary vaccines. The efficacy of alum in increasing antibody responses to diphtheria and tetanus toxoids is well established and an HBsAg vaccine has been adjuvanted with alum.

A wide range of extrinsic adjuvants can provoke potent immune responses to antigens. These include saponins complexed to membrane protein antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria and mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, and liposomes.

To efficiently induce humoral immune responses (HIR) and cell-mediated immunity (CMI), immunogens are often emulsified in adjuvants. Many adjuvants are toxic, inducing granulomas, acute and chronic inflammations (Freund's complete adjuvant, FCA), cytolysis (saponins and pluronic polymers) and pyrogenicity, arthritis and anterior uveitis (LPS and MDP). Although FCA is an excellent adjuvant and widely

used in research, it is not licensed for use in human or veterinary vaccines because of its toxicity.

Desirable characteristics of ideal adjuvants include:

(1) lack of toxicity;

(2) ability to stimulate a long-lasting immune response;

(3) simplicity of manufacture and stability in long-term storage;

(4) ability to elicit both CMI and HIR to antigens administered by various routes, if required;

(5) synergy with other adjuvants;

(6) capability of selectively interacting with populations of antigen presenting cells (APC);

(7) ability to specifically elicit appropriate T_H1 or T_H2 cell-specific immune responses; and

(8) ability to selectively increase appropriate antibody isotype levels (for example, IgA) against antigens.

U.S. Patent No. 4,855,283 granted to Lockhoff et al on August 8, 1989, which is incorporated herein by reference thereto, teaches glycolipid analogues including N-glycosylamides, N-glycosylureas and N-glycosylcarbamates, each of which is substituted in the sugar residue by an amino acid, as immuno-modulators or adjuvants. Thus, Lockhoff et al. 1991 (ref. 20) reported that N-glycolipid analogs displaying structural similarities to the naturally-occurring glycolipids, such as glycophospholipids and glycoglycerolipids, are capable of eliciting strong immune responses in both herpes simplex virus vaccine and pseudorabies virus vaccine. Some glycolipids have been synthesized from long chain-alkylamines and fatty acids that are linked directly with the sugars through the anomeric carbon

atom, to mimic the functions of the naturally occurring lipid residues.

U.S. Patent No. 4,258,029 granted to Moloney, assigned to the assignee hereof and incorporated herein by reference thereto, teaches that octadecyl tyrosine hydrochloride (OTH) functions as an adjuvant when complexed with tetanus toxoid and formalin inactivated type I, II and III poliomyelitis virus vaccine. Also, Nixon-George et al. 1990, (ref. 21) reported that octadecyl esters of aromatic amino acids complexed with a recombinant hepatitis B surface antigen, enhanced the host immune responses against hepatitis B virus.

2. Immunoassays

The lactoferrin receptor proteins, analogs and/or fragments thereof of the present invention are useful as immunogens, as antigens in immunoassays including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known in the art for the detection of anti-*Moraxella*, lactoferrin receptor protein antibodies. In ELISA assays, the lactoferrin receptor protein, analogs and/or fragments corresponding to portions of Lfr protein, are immobilized onto a selected surface, for example, a surface capable of binding proteins or peptides such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed lactoferrin receptor, analogs and/or fragments, a non-specific protein such as a solution of bovine serum albumin (BSA) or casein that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by non-specific bindings of antisera onto the surface.

The immobilizing surface is then contacted with a sample, such as clinical or biological materials, to be tested in a manner conducive to immune complex (antigen/antibody) formation. This procedure may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to incubate for from about 2 to 4 hours, at temperatures such as of the order of about 25° to 37°C. Following incubation, the sample-contacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween or a borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound lactoferrin receptor protein, analogs and/or fragments and subsequent washing, the occurrence, and even amount, of immunocomplex formation may be determined by subjecting the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and in general IgG. To provide detecting means, the second antibody may have an associated activity such as an enzymatic activity that will generate, for example, a color development upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for example, a spectrophotometer.

3. Use of Sequences as Hybridization Probes

The nucleotide sequences of the present invention, comprising the sequence of the lactoferrin receptor gene, now allow for the identification and cloning of the lactoferrin receptor genes from any species of *Moraxella*.

The nucleotide sequences comprising the sequence of the lactoferrin receptor genes of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other *lfr* genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other *lfr* genes. For a high degree of selectivity, relatively stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results. In general, convenient hybridization temperatures in the presence of 50% formamide are: 42°C for a probe which is 95 to 100% homologous to the target fragment, 37°C for 90 to 95% homology and 32°C for 85 to 90% homology.

Such hybridization conditions may be employed to determine DNA sequences which encode a functional lactoferrin receptor of *Moraxella* and which hybridize under stringent conditions to any one of the DNA sequences (a) or (b), described above.

In a clinical diagnostic embodiment, the nucleic acid sequences of the *lfr* genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide

variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin and digoxigenin-labelling, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with samples containing *lfr* gene sequences.

The nucleic acid sequences of *lfr* genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid-phase procedures, the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e. g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the *lfr* genes or fragments thereof of the present invention under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required depending on, for example, the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label. It is preferred to select nucleic acid sequence portions which are

conserved among species of *Moraxella*. The selected probe may be at least 18bp and may be in the range of about 30 to 90 bp.

4. Expression of the Lactoferrin Receptor Genes

Plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell may be used for the expression of the lactoferrin receptor genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the host cell for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM™-11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as *E. coli* LE392.

Promoters commonly used in recombinant DNA construction include the β-lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system as described in U.S. Patent No. 4,952,496. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with genes. The particular promoter used will generally be a matter of choice depending upon the desired results. Hosts that

are appropriate for expression of the lactoferrin receptor genes, fragments or analogs thereof, may include *E. coli*, *Bacillus* species, *Haemophilus*, fungi, yeast, *Moraxella*, *Bordetella*, or the baculovirus expression system may be used.

In accordance with this invention, it is preferred to produce the lactoferrin receptor protein, fragment or analog thereof, by recombinant methods, particularly since the naturally occurring LfR protein as purified from a culture of a species of *Moraxella* may include trace amounts of toxic materials or other contaminants. This problem can be avoided by using recombinantly produced LfR protein in heterologous systems which can be isolated from the host in a manner to minimize contaminants, including other proteins of the *Moraxella* strain, in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have LPS and are, therefore, endotoxin free. Such hosts include species of *Bacillus* and may be particularly useful for the production of non-pyrogenic lactoferrin receptor proteins, fragments or analogs thereof. Furthermore, recombinant methods of production permit the manufacture of Lbp1 or Lbp2 or ORF3 or respective analogs or fragments thereof, separate from one another which is distinct from the normal combined proteins present in *Moraxella*.

Sequence Alignment and Analysis

Sequence alignments were performed using the ALIGN (Trademark) or GENALIGN (Trademark) computer programs (Inteligenetics Suite 5.4, Oxford Molecular). ALIGN® uses the Needleman-Wunsch algorithm (ref. 35) and its later modifications to locate regions of similarity between two sequences. Finding regions of maximum

similarity between two sequences can be solved in a rigorous manner using the iterative matrix calculation of the Needleman and Wunsch 1997 algorithm. The analysis is restricted to regions with no internal deletions or insertions, joined by a minimum number of loop-outs or deletions. Sellers (ref. 36) developed a true metric measure of the "distance" between sequences and Waterman (ref. 37) extended this algorithm to include insertions and deletions of arbitrary length. Smith (ref. 38) improved the early algorithms to find the subsequences of maximum similarity. The algorithm has been used to analyze sequences as long as 5000 bases by dividing these sequences into segments of 200 to 400 bases, and then reassembling them into a final best match. This method of dividing the sequence and then reassembling it has proven quite robust. The algorithm permits the size of the segment to be specified which the program searches for similarities. The program then assembles the segments after checking overlaps of adjacent subsequences. The weighting of deletions and the relative size of overlaps may be controlled. The program displays the results to show the differences in closely related sequences.

GENALIGN® is a multiple alignment program. Up to 99 sequences using the Martinez/Regions (ref. 39) or Needleman-Wunsch (ref. 35) method may be analyzed for alignment. GENALIGN places the sequences in an order that puts the most closely aligned sequence pairs adjacent to each other. A consensus sequence is displayed under the multiple sequence alignments. The sequences used in developing the consensus sequence file for use in other programs. GENALIGN allows the parameters of the search to be changed so that alternate alignments of the sequences can be formed.

These programs are used employing their default settings. The default settings are as follows:

FastDB

AMINO-Res-length	=	2
DELetion-weight	=	5.00
LEngth-factor	=	0
Matching-weight	=	1.00
NUCLEIC-Res-length	=	4
SPread-factor	=	50

Findseq

Search Parameters:

Similarity matrix	Unitary
K-tuple	4
Mismatch penalty	1
Joining Penalty	30
Randomization group length	0
Cutoff score	5

Alignment Parameters:

Window size	32
Gap penalty	1.00
Gap size penalty	0.33

Such procedures may be used to determine DNA sequences which encode a functional lactoferrin receptor of *Moraxella* and which may have at least about 90% sequence identity with any one of the DNA sequences (a) or (b), described above.

Biological Deposits

Certain vectors that contain at least a portion coding for a lactoferrin receptor protein from strains of *Moraxella catarrhalis* strain 4223 and Q8 and a strain of *M. catarrhalis* RH408 that are described and referred to herein have been deposited with the American Type Culture Collection (ATCC) located at 12301 Parklawn Drive, Rockville, Maryland 20852, USA, pursuant to the

Budapest Treaty and prior to the filing of this application. Samples of the deposited vectors and bacterial strain will become available to the public and the restrictions imposed on access to the deposits will be removed upon grant of a patent based upon this United States patent application. In addition, the deposit will be replaced if viable samples cannot be dispensed by the Depository. The invention described and claimed herein is not to be limited in scope by the biological materials deposited, since the deposited embodiment is intended only as an illustration of the invention. Any equivalent or similar vectors or strains that encode similar or equivalent antigens as described in this application are within the scope of the invention.

Deposit Summary

Deposit	ATCC Designation	Date deposited
Plasmid pLD1-8	97,997	April 23, 1997
Plasmid pLDW1	97,998	April 23, 1997
Strain RH408	55,637	Dec. 9, 1994

EXAMPLES

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific Examples. These Examples are described solely for purposes of illustration and are not intended to limit the scope of the invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations.

Methods of molecular genetics, protein biochemistry and immunology used but not explicitly described in this disclosure and these Examples are amply reported in the scientific literature and are well within the ability of those skilled in the art.

Example 1

This Example illustrates the generation of oligonucleotide primers for PCR amplification of *M. catarrhalis lbpA*.

Native Lbp1 was purified by affinity chromatography using high stringency conditions as described in United States Patent Application No. 08/552,232, assigned to the assignee hereof and the disclosure of which is incorporated herein by reference, and in ref. 40.

The purified Lbp1 protein was digested overnight with cyanogen bromide, then fragments separated by SDS PAGE and submitted to sequence analysis on an AB1 model 477A. A 13 kDa protein fragment was found to have the N-terminal sequence MVQYTYRKGKENKAH (SEQ ID No: 22). A degenerate oligonucleotide primer (4393.RD) was prepared based upon this sequence:

Q Y T R K G E N K A (SEQ ID No: 28)

5' 3'

CAA TAT ACI CGT AAA GGT GAA AAT AAA GC (SEQ ID No: 29)
CAA TAT ACI CGT AAA GGC GAA AAC AAA GC (SEQ ID No: 30)
CAA TAT ACI CGT AAA GGT GAA AAC AAA GC (SEQ ID No: 31)
CAA TAT ACI CGT AAA GGC GAA AAT AAA GC (SEQ ID No: 32)
CAA TAT ACI CGC AAA GGC GAA AAC AAA GC (SEQ ID No: 33)
CAA TAT ACI CGC AAA GGC GAA AAT AAA GC (SEQ ID No: 34)
CAA TAT ACI CGC AAA GGT GAA AAT AAA GC (SEQ ID No: 35)
CAA TAT ACI CGC AAA GGT GAA AAC AAA GC (SEQ ID No: 36)

The Y^6 and K^{10} residues were omitted from the sequence analysis report for the N-terminal sequence and the oligonucleotides used to PCR amplify the 2.2 kb fragment were incorrect, but nevertheless were successful.

There is a conserved C-terminal pentapeptide found in all known Lbp1 and Tbp1 protein sequences: LEMKF (SEQ ID No. 26). An oligonucleotide primer (4572.RD) was prepared based upon the complementary DNA sequence encoding this pentapeptide:

L E M K F *		
5' CTT GAA ATG AAG TTT TAA 3'		(SEQ ID No: 37)
3' GAA CTT TAC TTC AAA ATT 5'	4572.RD	(SEQ ID No: 38)

Example 2

This Example illustrates the preparation of chromosomal DNA from *M. catarrhalis* strains 4223 and Q8.

M. catarrhalis isolate 4223 was inoculated into 100 ml of BHI broth, and incubated for 18 hr at 37°C with shaking. The cells were harvested by centrifugation at 10,000 x g for 20 min. The pellet was used for extraction of *M. catarrhalis* 4223 chromosomal DNA.

The cell pellet was resuspended in 20 ml of 10 MM Tris-HCl (pH 7.5)-1.0 mM EDTA (TE). Pronase and SDS were added to final concentrations of 500 µg/ml and 1.0%, respectively, and the suspension was incubated at 37°C for 2 hr. After several sequential extractions with phenol, phenol:chloroform (1:1), and chloroform:isoamyl alcohol (24:1), the aqueous extract was dialysed, at 4°C, against 1.0 M NaCl for 4 hr, and against TE (pH 7.5) for a further 48 hr with three buffer changes. Two volumes of ethanol were added to the dialysate, and the DNA was spooled onto a glass

rod. The DNA was allowed to air-dry, and was dissolved in 3.0 ml of water. Concentration was estimated, by UV spectrophotometry, to be about 290 µg/ml.

M. catarrhalis strain Q8 was grown in BHI broth. Cells were pelleted from 50 ml of culture by centrifugation at 5000 rpm for 20 minutes, at 4°C. The cell pellet was resuspended in 10 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) and proteinase K and SDS were added to final concentrations of 500 µg/ml and 1%, respectively. The sample was incubated at 37°C for 4 hours until a clear lysate was obtained. The lysate was extracted twice with Tris-saturated phenol/chloroform (1:1), and twice with chloroform. The final aqueous phase was dialysed for 24 hours against 2 X 1000 ml of 1 M NaCl at 4°C, changing the buffer once, and for 24 hours against 2 x 1000 ml of TE at 4°, changing the buffer once. The final dialysate was precipitated with two volume of 100% ethanol. The DNA was spooled, dried and resuspended in 5 to 10 ml of TE buffer.

Example 3

This Example illustrates the PCR amplification of a fragment of *M. catarrhalis lbpA* and the generation of probes for screening libraries.

PCR amplification was performed on chromosomal DNA isolated in Example 2 using primers 4393.RD and 4572.RD under the following cycling conditions: 25 cycles of 94°C for 1 min, 47°C for 30 sec and 72°C for 1 min. PCR4 is the amplification of the 4223 *lbpA* fragment and PCR5 is the amplification of the Q8 *lbpA* fragment. A specific band of about 2.2 kb was amplified and partial sequence analysis was performed to ensure that the gene product was related to *lbpA* and was not *tbpA*. The derived amino acid sequences are shown in Figure 1 and have been aligned with the complete 4223 Lbp1 sequence

to show their placement and the 4223 Tbp1 sequence (USAN 08/613,009) to indicate their uniqueness.

The full-length 2.2 kb gene fragment was randomly labeled with ^{32}P and used to probe genomic libraries.

Example 4

This Example illustrates the generation and screening of the EMBL 3 libraries.

Chromosomal DNA was prepared as described in Example 2. A series of *Sau*3AI restriction digests of chromosomal DNA, in final volumes of 10 µL each, were carried out in order to optimize the conditions necessary to generate maximal amounts of restriction fragments within a 15 to 23 kb size range. Using the optimized digestion conditions, a large-scale digestion was set up in a 100 µL volume, containing the following: 50 µL of chromosomal DNA (290 µg/ml), 33 µL water, 10 µL 10X *Sau*3A buffer (New England Biolabs), 1.0 µL BSA (10 mg/ml, New England Biolabs), and 6.3 µL *Sau*3A (0.04 U/µL). Following a 15 min. incubation at 37°C, the digestion was terminated by the addition of 10 µL of 100 mM Tris-HCl (pH 8.0)-10 mM EDTA-0.1% bromophenol blue-50% glycerol (loading buffer). Digested DNA was electrophoresed through a 0.5% agarose gel in 40 mM Tris acetate-2 mM $Na_2EDTA.2H_2O$ (pH 8.5) (TAE buffer) at 50 V for 6 hr. The region containing restriction fragments within a 15 to 23 kb molecular size range was excised from the gel, and placed into dialysis tubing containing 3.0 ml of TAE buffer. DNA was electroeluted from the gel fragment by applying a field strength of 1.0 V/cm for 18 hr. Electroeluted DNA was extracted once each with phenol and phenol:chloroform (1:1), and precipitated with ethanol. The dried DNA was dissolved in 5.0 µL water.

Size-fractionated chromosomal DNA was ligated with

*Bam*HI-digested EMBL3 arms (Promega), using T4 DNA ligase in a final volume of 9 µL. The entire ligation mixture was packaged into lambda phage using a commercial packaging kit (Amersham), following manufacturer's instructions.

The packaged DNA library was amplified on solid media. 0.1 ml aliquots of *Escherichia coli* strain NM539 in 10 mM $MgSO_4$ (OD_{260} = 0.5) were incubated at 37°C for 15 min. with 15 to 25 µL of the packaged DNA library. Samples were mixed with 3 ml of 0.6% agarose containing 1.0% BBL trypticase peptone-0.5% NaCl (BBL top agarose), and mixtures were plated onto 1.5% agar plates containing 1.0% BBL trypticase peptone-0.5% NaCl, and incubated at 37°C for 18 hr. 3 ml quantities of 50 mM Tris-HCl (pH 7.5)-8 mM magnesium sulfate heptahydrate-100 mM NaCl-0.01% (w/v) gelatin (SM buffer) were added to each plate, and plates were left at 4°C for 7 hr. SM buffer containing phage was collected from the plates, pooled together, and stored in a screwcap tube at 4°C, with chloroform.

Ten µL aliquots of phage stock were combined each with 100 µL of *E. coli* strain LE392 in 10 mM MgSO4 (OD_{260} = 0.5)(plating cells), and incubated at 37°C for 15 min. The samples were mixed with 3 ml each of BBL top agarose, and the mixtures were poured onto 1.5% agarose plates containing 1% bacto tryptone-0.5% bacto yeast extract-0.05% NaCl (LB agarose; Difco) and supplemented with 200 µM EDDA. The plates were incubated at 37°C for 18 hr. Plaques were lifted onto nitrocellulose filters (Amersham Hybond-C Extra) which were hybridized with the 32P-labelled 2.2 kb PCR fragment. Several putative phage clones were obtained from each library and clones 4223LfR.17 and Q8LfR.13 were chosen for further analysis.

Example 5

This Example illustrates the subcloning of the phage clones containing *M. catarrhalis lfr* genes.

Restriction enzyme analysis and Southern blotting using the screening probes, indicated that at least a portion of *lbpA* was localized to an about 9 kb Hind III fragment from each phage clone. The about 9 kb Hind III fragment from 4223LfR.17 was subcloned into pUC 18, generating clone pLD1-8. The about 9 kb Hind III fragment from Q8LfR.13 was subcloned into pBluescript, generating plasmid pLDW1. Internal about 5.5 kb EcoR V fragments were subcloned generating plasmids pLD3 and pLDW3 for the 4223 and Q8 genes, respectively.

Example 6

This Example illustrates the sequence analysis of clones containing the *M. catarrhalis lfr* genes from strains 4223 and Q8.

Sequence analysis of the 5.5 kb EcoR V fragments from pLD3 and pLDW3, revealed that they each contained the 3'-end of *lbpB*, the complete *lbpA* gene, and a third complete gene designated *orf3*. The remainder of the *lbpB* genes was found on the about 9 kb Hind III fragments from pLD1-8 and pLDW1. Partial restriction enzyme analysis of the 4223 *lbpA*, *lbpB*, and *orf3* genes, based upon the nucleotide sequences is shown in Figure 3. Partial restriction enzyme analysis of the Q8 *lbpA*, *lbpB*, and *orf3* genes, based upon the nucleotide sequences is shown in Figure 5. The complete sequences of the *lbpB*, *lbpA*, and *orf3* genes comprising the putative *lfr* locus from *M. catarrhalis* 4223 and Q8 is shown in Figures 2 and 4, respectively. The intergenic distance between the *lbpB* and *lbpA* genes is 184 nucleotides, while a single nucleotide separates the

lbpA and *orf3* genes. A putative promoter and ribosome binding site is indicated by underlining upstream of both *lbpB* and *lbpA*. A fourth potential gene was cloned on the approximately 9 kb Hind III fragments.

The N-terminal sequence of the native Lbp1 protein is unknown. Examination of the deduced amino acid sequence of the *lbpA* gene indicates that there are two possible ATG start codons at positions 1 and 16. The first position is downstream of strong promoter elements found in the *lbpB-lbpA* intergenic region and the second position is followed by a putative signal sequence. The *M. catarrhalis* 4223 and Q8 Lbp1 proteins (from the first ATG) have molecular mass values of about 110 kDa and are 99% identical. The deduced Lbp1 protein sequences from *M. catarrhalis* strains 4223 and Q8 are compared in Figure 6. They are also compared with the *iroA/lbpA* gene from *N. meningitidis* strain BNCV (ref. 24) and the *lbpA* gene from *N. gonorrhoeae* strain FA19 (ref. 25). The *M. catarrhalis* proteins are found to be about 32% identical and about 50% similar to the *Neisseria* proteins. As shown in Figure 1, there is very limited sequence homology between the *M. catarrhalis* Tbp1 and Lbp1 sequences.

The deduced Lbp2 protein sequences from *M. catarrhalis* strains 4223 and Q8 are compared in Figure 7. The 4223 and Q8 Lbp2 proteins both have molecular masses of about 99 kDa and are 92% identical and 95% similar to each other. A comparison to the *M. catarrhalis* Tbp2 proteins shows very little homology except the LEGGFY (SEQ ID No: 27) epitope previously identified in *H. influenzae* and *N. meningitidis* Tbp2 proteins (Fig. 8). A cysteine residue at position 32 is preceded by a consensus sequence for lipoproteins suggesting that Lbp2, like Tbp2, is a lipoprotein. An

unusual feature of the Lbp2 proteins is the high combined aspartic acid and asparagine content which is nearly 20%. In addition, the 4223 Lbp2 amino acid composition from residues 698 to 751 is about 52% aspartic acid.

The 4223 and Q8 *lfr orf3* genes would encode proteins of molecular mass about 60 kDa, respectively. A notable feature of the ORF3 protein is a potential signal sequence, a terminal phenylalanine which is often associated with membrane anchored proteins, an internal repeat sequence of DGLG (SEQ ID No: 39), and a high leucine content of 15%. The deduced Lbp3 protein sequences are compared in Figure 9. These proteins are 98% identical and 99% similar.

Example 7

This Example illustrates the construction of vectors to express *M. catarrhalis* Lbp1 from the first methionine in *E. coli*.

There are two possible start codons at the beginning of the *lbpA* gene and hence two expression constructs were made. The construction scheme for 4223 or Q8 *lbpA* expressed from the first methionine is shown in Figure 10. An approximately 200 bp fragment of the 5'-end of *lbpA* from the ATG to a BstE II site was PCR amplified using primers 5405.RD and 5407.RD. An Nde I site was engineered at the 5'-end to facilitate cloning into the pT7-7 vector.

NdeI

M S K S I T (SEQ ID No: 40)

5' GGAATTCCAT ATG TCA AAA TCT ATC ACA AA 3' 5405.RD (SEQ ID No: 41)

BstE II

L D A I T V T A A (SEQ ID No: 42)

5' T TTA GAT GCC ATC ACG GTA ACC GCC GCC CC 3' (SEQ ID No: 43)

3' A AAT CTA CGG TAG TGC CAT TGG CGG CGG GG 5' 5407.RD
(SEQ ID No: 44)

In order to subclone the *lbpA* gene into pT7-7, a approximately 515 bp fragment of the 3'-end of the gene from an Sph I site to the stop codon was PCR amplified using primers 5281.RD and 5282.RD and a BamH1 site was engineered at 3'-end.

Sph I

G K L D L H A M T S (SEQ ID No: 45)
5' GGC AAA CTG GAT TTG CAT GCC ATG ACA TCA 3' 5281.RD
(SEQ ID No: 46)

S L E M K F * (SEQ ID No: 47)
5' AGT CTT GAA ATG AAG TTT TAA 3'
(SEQ ID No: 48)
3' TCA GAA CTT TAC TTC AAA ATT GCC CTA GGG C 5' 5282.RD
BamH I (SEQ ID No: 49)

For the Q8 subclone, plasmid pLDW3, prepared as described in Example 5, was digested with BstE II and Sph I generating a 2.3kb fragment of *lbpA* which was ligated with the Nde I-BstE II and SphI-BamH I PCR fragments and cloned into pT7-7 digested with NdeI and BamH I. The resulting plasmid pQW1A thus contains the full-length Q8 *lbpA* gene from the first methionine, under the control of the T7 promoter. DNA from pQW1A was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain QW1A which was grown and induced using IPTG. Expressed proteins were resolved by SDS-PAGE and the induced Lbp1 protein was visualized by Coomassie blue staining (Fig. 11).

For the 4223 subclone, plasmid pLD3, prepared as described in Example 5 was digested with BstEII and SphI, generating a 2.3 kb fragment of *lbpA*, which was

ligated with the Nde I-BstE II and SphI-BamH I PCR fragments and cloned into pT7-7 digested with NdeI and BamH I. The resulting plasmid pRD1A thus contains the full-length 4223 *lbpA* gene from the first possible methionine under the control of the T7 promoter. DNA from pRD1A was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate strain RD1A which was grown and induced using IPTG. Expressed proteins were resolved by SDS-PAGE and the induced Lbp1 protein was visualized by Coomassie blue staining (Fig. 11).

The Q8 Lbp1 protein was expressed at very high levels but the 4223 Lbp1 protein was expressed at substantially lower levels.

Example 8

This Example illustrates the extraction and purification of rLbp1 from *E. coli*. The procedure is illustrated generally in Figure 14.

E. coli cells from a 500 ml culture, prepared as described in Example 7, were resuspended in 40 ml of 50 mM Tris-HCl, pH 8.0 containing 5 mM AEBSF (protease inhibitor) and 0.1 M NaCl, and disrupted by sonication (3 x 10 min, 70% duty circle). The extract was centrifuged at 20,000 x *g* for 30 min and the resultant supernatant, which contained greater than 95% of the soluble proteins from *E. coli*, was discarded. The remaining pellet (Figure 14, PPT1) was further extracted in 40 ml of 50 mM Tris, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA. The mixture was stirred at 4°C for at least 1 hour and then centrifuged at 20,000 x *g* for 30 min and the supernatant containing residual soluble proteins and the majority of the membrane proteins was discarded. The resultant pellet (Figure 14, PPT2) was further extracted in 40 ml of 50

mM Tris, pH 8.0 containing 1% octylglucoside. The mixture was stirred at 4°C for at least 1 hour and then centrifuged at 20,000 x *g* for 30 min. The supernatant containing residual contaminating proteins was discarded. The resultant pellet (Figure 14, PPT3) obtained after the above extractions contained the Lbp1 protein as inclusion bodies.

The rLbp1 protein was solubilized from the inclusion bodies in 50 mM Tris, pH 8.0, containing 6 M guanidine and 5 mM DTT. After centrifugation, the resultant supernatant was further purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris-HCl, pH 8.0, containing 2 M guanidine and 5 mM DTT. The fractions were analysed by SDS-PAGE and those containing purified rLbp1 were pooled. Triton X-100 was added to the pooled rLbp1 fraction to a final concentration of 0.1%. The fraction was dialysed overnight at 4°C against PBS, and then centrifuged at 20,000 x *g* for 30 min. The purified rLbp1 was stored at -20°C. Samples from the purification were analyzed by SDS-PAGE (Fig. 15).

Example 9

This Example illustrates the construction of vectors to express *M. catarrhalis* Lbp1 from the second methionine in *E. coli*.

The construction scheme for 4223 or Q8 *lbpA* expressed from the second methionine is shown in Figure 10. An approximately 200 bp fragment of the 5'-end of *lbpA* from the ATG to a BstE II site was PCR amplified using primers 5406.RD and 5407.RD. An Nde I site was engineered at the 5'-end to facilitate cloning into the pT7-7 vector.

NdeI

M T T H R L (SEQ ID No: 50)

5' GGAATTCCAT ATG ACC ACG CAC CGC TTA AA 3' 5406.RD

(SEQ ID No: 51)

BstE II

L D A I T V T A A

5' T TTA GAT GCC ATC ACG GTA ACC GCC GCC CC 3'

3' A AAT CTA CGG TAG TGC CAT TGG CGG CGG GG 5' 5407.RD

The 3'-end of the *lbpA* gene was PCR amplified from the SphI restriction site to the stop codon using primers 5281.RD and 5282.RD as described in Example 8. The 2.3 kb BstE II-Sph I fragments described in Example 8 were ligated to the Nde I-BstE II and Sph I-BamH I PCR fragments and cloned into pT7-7 that had been digested with NdeI and BamH I. Plasmid pQW1B thus contains a full-length Q8 *lbpA* gene from the second methionine and plasmid pRD1B contains a full-length 4223 *lbpA* gene from the second methionine under the direction of the T7 promoter. DNA was purified and transformed by electroporation into electrocompetent BL21(DE3) cells to generate recombinant strains which were grown and induced using IPTG. Expressed proteins were resolved by SDS-PAGE and the induced Lbp1 proteins were visible by Coomassie blue staining (Fig. 11).

As seen for the longer protein in Example 8, the shorter Lbp1 from Q8 was expressed to much higher levels than the corresponding 4223 protein.

Example 10

This Example illustrates the construction of vectors to express *M. catarrhalis* Lbp2 with a leader sequence from *E. coli*.

The construction scheme is illustrated in Figure 12. There are two BspH I sites within the *lbpB* genes of strains 4223 and Q8. The 5'-end of the *lbpB* gene was PCR amplified from the ATG start codon through the

first BspH I site generating an approximately 201 bp fragment. An NdeI site was engineered at the ATG to facilitate cloning into the pT7-7 expression vector. The oligonucleotides used for amplification are illustrated below:

NdeI

M S T V K T P H (SEQ ID No: 52)

5' GGAATTCCAT ATG AGT ACT GTC AAA ACC CCC CAC A 3' 5533.RD

(SEQ ID No: 53)

BspH I

I P N T G H D N T N (SEQ ID No: 54)

5' A ATA CCG AAC ACA GGT CAT GAC AAC ACC AAT 3'

(SEQ ID No: 55)

T TAT GGC TTG TGT CCA GTA CTG TTG TGG TTA 5' 5534.RD

(SEQ ID No: 56)

The 3'-end of the *lbpB* gene was PCR amplified from the second BspH I site to the TAA stop codon generating a 381 bp fragment. A BamH I site was introduced after the stop codon for cloning purposes. The oligonucleotides used for amplification are illustrated below:

N E P T H E K T F A (SEQ ID No: 57)

5' AAT GAG CCT ACT CAT GAA AAA ACC TTT GCC 3' 5535.RD

(SEQ ID No: 58)

G A V F G A V K D K * (SEQ ID No: 59)

5' GG GCT GTC TTT GGG GCT GTT AAA GAT AAA TAA 3'

(SEQ ID No: 60)

CC CGA CAG AAA CCC CGA CAA TTT CTA TTT ATT CCTAGGGC 5' 5536.RD

BamH I (SEQ ID No. 61)

Plasmids pLD1-8 or pLDW1, prepared as described in Example 4, were digested with BspH I to release a 2.1 kb internal fragment of the *lbpB* gene which was ligated with the 5'- and 3'-PCR fragments and cloned into pT7-7

that had been digested with NdeI and BamH I. The resulting plasmids, pLD2A and pLDW2A, contain the full-length 4223 and Q8 *lbpB* genes under the control of the T7 promoter, respectively.

Example 11

This Example illustrates the construction of vectors to express the mature *M. catarrhalis* Lbp2 proteins from *E. coli*.

The construction scheme is illustrated in Figure 12. The putative mature Lbp2 lipoproteins start at the Cys^{32} residue. A scheme similar to that described in Example 10 can be used to generate expression clones. To amplify the 5'-end of the *lbpB* gene, a sense PCR primer is designed that includes an NdeI site for subsequent cloning and an ATG start codon for initiation of translation followed immediately by the Cys^{32} residue. The antisense primer is the same as that described in Example 9 (5534.RD) and includes the BspH I cloning site. The amplified fragment is ~112 bp long. The oligonucleotides are illustrated below:

NdeI

M C R S D D I S V N (SEQ ID No: 62)

5' GGAATTCCAT ATG TGC CGC TCT GAT GAC ATC AGC GTC AAT 3' .RD (SEQ ID No: 63)

BspH I

I P N T G H D N T N (SEQ ID No: 54)

5' A ATA CCG AAC ACA GGT CAT GAC AAC ACC AAT 3' (SEQ ID No: 55)

3' T TAT GGC TTG TGT CCA GTA CTG TTG TGG TTA 5' 5534.RD (SEQ ID No: 56)

The BspH I-BamH I 3'-end of the *lbpB* gene is PCR amplified as in Example 9 and the plasmid expressing

mature Lbp2 is constructed by ligating the 5'- and 3'-PCR fragments with the 2.1 kb BspH I fragment and vector pT7-7 digested with NdeI and BamH I. The resulting plasmids, pLD2B and pLDW2B, contain the *lbpB* gene encoding the mature Lbp2 proteins from strains 4223 and Q8 under the direction of the T7 promoter, respectively.

Example 12

This Example illustrates the construction of a vector to express the *M. catarrhalis lfr* Lbp3 from *E. coli*.

The construction scheme is illustrated in Figure 13. Oligonucleotides were used to generate the 5'-end of the *orf3* gene from the ATG start codon to an AlwN I site. An NdeI site was engineered at the 5'-end for subsequent cloning into pT7-7. The oligonucleotides are shown below:

```
    NdeI
      M   T   C   L   P   K   T   N   P   A   L   K   V   K   H   R
5'  T ATG ACC TGT TTA CCA AAG ACC AAC CCT GCT TTA AAA GTC AAG CAC AGA
3'    AC TGG ACA AAT GGT TTC TGG TTG GGA CGA AAT TTT CAG TTC GTG TCT
                 AlwN I
```

F L K Q V				(SEQ ID No: 64)
TTT TTA AAG CAG GTG	3'	5532.RD		(SEQ ID No: 65)
AAA AAT TTC GTC	5'	5457.RD		(SEQ ID No: 66)

The pLD1-8 or pLDW1 plasmid, prepared as described in Example 5, was digested with BstE II generating a 4.6 kb fragment which was filled in with Klenow polymerase before being digested with AlwNI. The resultant 1.8 kb fragment was ligated with the annealed NdeI-AlwN I oligonucleotides and cloned into pT7-7 that had been digested with NdeI and SmaI. The resulting plasmids, pLRD3 and pLQW3, contain the full-length *orf3*

genes from strains 4223 and Q8 under the direction of the T7 promoter, respectively.

Example 13

This Example describes the cloning and sequencing of the *lbpB* gene from *M. catarrhalis* strain VH19.

Chromosomal DNA was prepared from *M. catarrhalis* strain VH19, as described previously in Example 2. Oligonucleotide primers were designed based upon the flanking sequence of the 4223 *lbpB* gene. The sense primer was 5' AAGCTTAGCATGATGGCATCGGCT 3' (SEQ ID No: 67) and the antisense primer was 5' TTAGCCCAAGGCAAATCTGGTGCA 3' (SEQ ID No: 68). PCR was performed in buffer containing 10mM Tris-HCl (pH 8.3), 50 mM potassium chloride and 1.5 mM magnesium chloride. Each 100 µl reaction mixture contained 1 µg chromosomal DNA, 0.1 µg each primer, 2.5 units amplitaq DNA polymerase (Perkin Elmer Cetus, Foster City, California) and 10 mM of each dNTP (Perkin Elmer Cetus). The cycling conditions were 24 cycles of 94°C for 1 min, 47°C for 30 sec and 72°C for 1 min. Specific 2.9 kb fragments were amplified from two independent reactions and subcloned into pCR II (Invitrogen, Carlsbad, California), generating plasmids pVH19pcr1 and pVH19pcr2 for sequence analysis. A third PCR amplification was performed without subcloning the resultant DNA. Plasmid DNA from pVH19pcr1 and pVH19pcr2 was prepared from 50 ml overnight cultures using the Qiagen Plasmid Midi kit (Qiagen Inc, Chatsworth, California). PCR amplified DNA was purified for direct sequencing using a Qiagen PCR purification kit. DNA samples were sequenced on an ABI model 373A DNA sequencer using dye terminator chemistry. Oligonucleotide primers 17 to 25 bases in length were used to sequence both strands of the DNA.

The nucleotide sequence (SEQ ID No: 69) of the VH19 *lbpB* gene and the deduced amino acid sequence of the corresponding Lbp2 protein (SEQ ID No: 70) are shown in Figure 16. The encoded VH19 Lbp2 protein is 906 amino acids and is 77% identical and 84% similar to the 4223 and Q8 Lbp2 proteins. There is a putative lipoprotein signal sequence which is very similar to the 4223 and Q8 signal sequences. The high Asp and Asn content found in the 4223 and Q8 Lbp2 proteins is also present in the VH19 LbpB protein, as is the RGD sequence. A partial restriction map of the VH19 *lbpB* gene is shown in Figure 17.

An alignment of the Lbp2 proteins from *M. catarrhalis* strains 4223, Q8 and VH19 is shown in Figure 7. The *M. catarrhalis* Lbp2 proteins are also compared with partial Lbp2 sequences from *N. meningitis* strains BNCV (ref. 31) and H44/76 (ref. 24) and *N. gonorrhoeae* strain FA19 (ref. 25). There are small scattered regions of sequence homology to the known bacterial Tbp2 proteins (ref. 32). Residues that are conserved among the Tbp2 proteins and the *M. catarrhalis* Lbp2 proteins are underlined in Figure 7 and include the LEGGFYG (SEQ ID No: 75) motif.

Example 14

This Example describes the construction of vectors for expression of the *M. catarrhalis* Lbp2 protein.

By analogy with Tbp2 proteins, Lbp2 was assumed to be a lipoprotein and constructs were designed for expression of Lbp2 with or without a lipopeptide signal sequence. There is a unique Bgl I site in *lbpb*. To express the full-length Lbp2 protein with leader sequence (construct A), an approximately 429 bp 5'-fragment from the Met^1 start codon to the Bgl I site was PCR amplified and to express the mature protein

(construct B), an approximately 329 bp 5'-fragment from the putative Cys^{32} start to the Bgl I site was PCR amplified. The following sense primers were used:

Nde I

M S T V K T P H (SEQ ID No: 52)

5' GGAATTCCAT ATG AGT ACT GTC AAA ACC CCC CAC A 3' (SEQ ID No: 53)

for construct A or

Nde I

M C R S D D I S V N (SEQ ID No: 62)

5' GGAATTCCAT ATG TGC CGC TCT GAT GAC ATC AGC GTC AAT 3' (SEQ ID No: 63)

for construct B and the anti-sense primer was:

G K N L R G P I (SEQ ID No: 72)

5' GGT AAA AAC TTG CGT CAG CCC ATC 3' (SEQ ID No: 73)

3' CCA TTT TTG AAC GCA GTC GGG TAG 5' (SEQ ID No: 74)

Bgl I

The Q8 *lfr*-containing plasmid, pLDW1 (Example 5), was digested with Bgl I and EcoR I to release a 2.3 kb *lbpB* fragment which was ligated with the Nde I - Bgl I PCR fragment and cloned into pT7-7 that had been digested with Nde I and EcoR I. The resulting plasmids, pQW2A and pQW2B, thus contain the Q8 *lbpB* gene encoding the full-length or mature Lbp2 proteins under the direction of the T7 promoter. The plasmids expressing the 4223 full-length or mature Lbp2 proteins were constructed in a similar manner and designated pRD2A and pRD2B. There was no measurable expression of rLbp2 from constructs containing the signal sequence, however the mature rLbp2 proteins were expressed at 5 to 10% of total proteins as inclusion bodies and were purified by the same process as that described for rLbp1 in Example 8. Samples from the purification were analyzed by SDS-PAGE (Figure 18).

Example 15

This Example describes the functional characterization of the recombinant lactoferrin binding proteins.

Human lactoferrin (Sigma) was conjugated to horseradish peroxidase using an EZ-Link maleimide activated horseradish peroxidase (HRP) kit (Pierce, Rockford, Illinois) according to the manufacturer's instructions. The lactoferrin binding activity of rLbp1 or rLbp2 was assessed by modifying the procedure described for transferrin binding proteins (ref. 17). Briefly, purified rLbp1 or rLbp2 was subjected to discontinuous electrophoresis through a 12.5% SDS PAGE gel. The proteins were electrophoretically transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore, Bedford, Massachusetts) and incubated with horseradish peroxidase-conjugated human lactoferrin (1:20 dilution) at 4°C overnight. LumiGLO substrate (Kirkegaard and Perry Laboratories, Inc., Gaithersburg, Maryland) was used for chemiluminescent detection of HRP activity according to the manufacturer's instructions. The Q8 rLbp1 protein did not bind human lactoferrin under these conditions, but the 4223 rLbp2 and Q8 rLbp2 proteins did (Fig. 19).

Example 16

This Example describes the immunization of animals and immunoassays.

Groups of two guinea pigs (Hartley outbred, Charles River, Quebec) were immunized intramuscularly (i.m.) with 5 µg doses of purified rLbp1 or rLbp2 protein emulsified in CFA or IFA. Anti-Lbp antibody titers in guinea pig immune sera were determined by antigen-specific ELISA. Microtiter wells (Nunc-MAXISORB, Nunc, Denmark) were coated with 50 µl of protein (0.5 µg ml^{-}

[1]). The reactive titer of an antiserum was defined as the reciprocal of the highest dilution consistently showing a two-fold increase in absorbance at 450 nm over that obtained with the pre-immune serum samples. The recombinant proteins elicited high titre antibodies as shown in Tables 1 and 2.

Example 17

This Example describes the antigenic conservation of Lbp1 and Lbp2 in *M. catarrhalis* strains.

To demonstrate the iron-dependent expression of the *lbpA* and *lbpB* genes, representative *M. catarrhalis* strains were grown in BHI ± 25 mM EDDA. Whole cell lysates were separated by SDS PAGE and electrophoretically transferred to nitrocellulose membrane. Guinea pig anti-Q8 rLbp1, anti-Q8 rLbp2 and anti-4223 rLbp2 antisera were used as first antibodies and horseradish peroxidase-conjugated protein G (ZYMED) was used as secondary antibody. To assess antigenic conservation, approximately 90 *M. catarrhalis* strains, obtained from North America or Finland were grown in BHI + 25 mM EDDA, and immunoblots were probed with guinea pig anti-4223 rLbp2 antibody, as above. All strains showed a protein band reactive with anti-rLbp2 antibody. There was very little size heterogeneity for the Lbp2 proteins from the 90 *M. catarrhalis* strains, ranging from approximately 100 kDa to 105 kDa. Representative immunoblots are illustrated in Fig. 19.

Example 18

This Example describes the assay used to determine the bactericidal antibody activity of anti-Lbp antibodies.

The assay was performed as described by ref. 33. Briefly, the *M. catarrhalis* strains were grown to an OD_{578} of 0.5 in BHI medium containing 25 mM EDDA. The

bacteria were diluted so that the pre-bleed control plates contained 100 to 300 cfu. Guinea pig anti-rLbp1 or anti-rLbp2 antisera and pre-bleed controls, were heated to 56°C for 30 min to inactivate endogenous complement and were diluted 1:64 with veronal buffer containing 0.1% BSA (VBS). Guinea pig complement (Biowhittaker, Walkersville, Maryland) was diluted 1:10 in VBS. Twenty-five µl each of diluted antiserum, bacteria and complement were added to duplicate wells of a 96 well microtiter plate (Nunc). The plates were incubated at 37°C for 60 min, gently shaking at 70 rpm on a rotary platform. Fifty µl of each reaction mixture were plated onto Mueller Hinton agar plates (Becton-Dickinson, Cockeysville, Maryland) which were incubated at 37°C for 24 h, then room temperature for 24 h, before the bacteria were counted. Antisera were determined to be bactericidal if ≥ 50% of bacteria were killed compared with negative controls.

Six strains of different geographical and anatomical origins were tested. The data in Table 3 illustrates that anti-4223 rLbp2 antibody was bactericidal for the homologous strain and three of five heterologous strains.

SUMMARY OF THE DISCLOSURE

In summary of this disclosure, the present invention provides purified and isolated DNA molecules containing lactoferrin receptor genes from *Moraxella catarrhalis*, the sequences of these lactoferrin receptor genes, and the derived amino acid sequences thereof. The genes and DNA sequences are useful for diagnosis, immunization, and the generation of diagnostic and immunological reagents. Immunogenic compositions, including vaccines, based upon expressed recombinant Lbp1 and/or Lbp2 and/or ORF3, portions

thereof, or analogs thereof, can be prepared for prevention of diseases caused by *Moraxella*. Modifications are possible within the scope of this invention.

Table 1

Bactericidal antibody titres for anti-native Lbp1

	Bactericidal titre - RH408		Bactericidal titre - Q8	
Antibody	Pre-immune	Immune	Pre-immune	Immune
Anti-4223 Lbp1	<8	114-330	<8	128-512

Bactericidal titres are expressed as the reciprocal dilution of antiserum capable of killing 50% of *M. catarrhalis* cells

Table 2 ELISA titers for guinea pig anti-Lbp antibodies raised against recombinant lactoferrin binding proteins

Coated antigen	Anti-Q8 rLbp1	Anti-Q8 rLbp2	Anti-4223 rLbp2
Q8 rLbp1	3,200 25,600	-	-
Q8 rLbp2	-	1,638,400 1,638,400	409,600 409,600
4223 rLbp2	-	409,600 409,600	819,200 819,200

Table 3. Bactericidal antibody activity of guinea pig anti-rLbp2 antibodies

Strain	locale[1]	source[2]	Lbp2 size	Bactericidal antibody activity[3]	
				Anti-4223 rLbp2	Anti-Q8 rLbp2
4223	New York	MEF	105 kDa	++	-
Q8	Quebec	sputum	105 kDa	±	-
VH19	Texas	MEF	105 kDa	+	NT[4]
LES-1	Finland	MEF	102 kDa	-	NT
H-04	Nova Scotia	MEF	100 kDa	+	NT
3	New York	sputum	100 kDa	++	NT

[1] geographic locale where strain was isolated
[2] anatomical source of clinical isolate. MEF is middle ear fluid from otitis media patients
[3] killing by antiserum diluted 1:64, compared to negative controls: - indicates 0-25% killing; ± indicates 26-49% killing; + indicates 50-75% killing; ++ indicates 76-100% killing.
[4] NT = not tested

REFERENCES

1. Brorson, J-E., A. Axelsson, and S.E. Holm. 1976. Studies on Branhamella catarrhalis (Neisseria catarrhalis) with special reference to maxillary sinusitis. Scan. J. Infect. Dis. 8:151-155.

2. Catlin, B.W., 1990. Branhamella catarrhalis: an organism gaining respect as a pathogen. Clin. Microbiol. Rev. 3: 293-320.

3. Hager, H., A. Verghese, S. Alvarez, and S.L. Berk. 1987. Branhamella catarrhalis respiratory infections. Rev. Infect. Dis. 9:1140-1149.

4. McLeod, D.T., F. Ahmad, M.J. Croughan, and M.A. Calder. 1986. Bronchopulmonary infection due to M. catarrhalis. Clinical features and therapeutic response. Drugs 31(Suppl.3):109-112.

5. Nicotra, B., M. Rivera, J.I. Luman, and R.J. Wallace. 1986. Branhamella catarrhalis as a lower respiratory tract pathogen in patients with chronic lung disease. Arch.Intern.Med. 146:890-893.

6. Ninane, G., J. Joly, and M. Kraytman. 1978. Bronchopulmonary infection due to Branhamella catarrhalis 11 cases assessed by transtracheal puncture. Br.Med.Jr. 1:276-278.

7. Srinivasan, G., M.J. Raff, W.C. Templeton, S.J. Givens, R.C. Graves, and J.C. Mel. 1981. Branhamella catarrhalis pneumonia. Report of two cases and review of the literature. Am. Rev. Respir. Dis. 123:553-555.

8. West, M., S.L. Berk, and J.K. Smith. 1982. Branhamella catarrhalis pneumonia., South. Med. J. 75:1021-1023.

9. Christensen, J.J., and B. Bruun. 1985. Bacteremia caused by a beta-lactamase producing strain of Branhamella catarrhalis. Acta. Pathol. Microbiol. Immunol. Scand. Sect. B 93:273-275.

10. Craig, D.B., and P.A. Wehrle. 1983. Branhamella catarrhalis septic arthritis. J. Rheumatol. 10:985-986.

11. Guthrie, R., K. Bakenhaster, R. Nelson, and R. Woskobnick. 1988. Branhamella catarrhalis sepsis: a case report and review of the literature. J. Infect. Dis. 158:907-908.

12. Hiroshi, S., E.J. Anaissie, N. Khardori, and G.P. Bodey. 1988. Branhamella catarrhalis septicemia in patients with leukemia. Cancer 61:2315-2317.

13. O'Neill, J.H., and P.W. Mathieson. 1987. Meningitis due to Branhamella catarrhalis. Aust. N.Z. J. Med. 17:241-242.

14. Murphy, T.F. 1989. The surface of Branhamella catarrhalis: a systematic approach to the surface antigens of an emerging pathogen. Pediatr. Infect. Dis. J. 8:S75-S77.

15. Van Hare, G.F., P.A. Shurin, C.D. Marchant, N.A. Cartelli, C.E.Johnson, D. Fulton, S. Carlin, and C.H. Kim. Acute otitis media caused by Branhamella catarrhalis: biology and therapy. Rev. Infect. Dis. 9:16-27.

16. Jorgensen, J.H., Doern, G.V., Maher, L.A., Howell, A.W., and Redding, J.S., 1990 Antimicrobial resistance among respiratory isolates of Haemophilus influenza, Moraxella catarrhalis, and Streptococcus pneumoniae in the United States. Antibicrob. Agents Chemother. 34: 2075-2080.

17. Schryvers, A.B. and Lee, B.C. (1988) Comparative analysis of the transferrin and lactoferrin binding proteins in the family Neisseriaceae. Can. J. Microbiol. 35, 409-415.

18. O'Hagan, DT. 1992. Oral deleivery of vaccines. Formulation and clinical pharmaco kinetic considerations. Clin. Pharmacokinet 22(t): 1-10.

19. Ulmer et al. 1993. Curr. Opinion Invest. Drugs 2:983-989.

20. Lockhoff, O., 1991. Glycolipids as immunomodulators: Synthesis and properties.

21. Nixon-George A., et al., 1990. The adjuvant effect of stearyl tyrosine on a recombinant subunit hepatitis B surface antigen. J Immunol 144 (12): 4798-4802.

22. Wallace, R.J. et al., 1990. Antibiotic susceptibilites and drug resistance in Moraxella (Branhaemella) catarrhalis. Am. J. Med. 88(5A): 465-505.

23. Nissinen A, et al., 1995. Development of beta-lactamase-mediated resistance to penicillin in middle-ear isolates of Moraxella catarrhalis in

Finnish children, 1978-1993. Clin Infect Dis 21 (5): 1193-1196.

24. Pettersson, A., et al., 1994. Identification of iroa Gene Product of Neisseria meningitides as a Lactoferrin Receptor. J. Bacteriol. 176(6): 1764-1766.

25. Biswas GD, Sparring PF. 1995. Characterization of lbpa, the structural gene for a lactoferrin receptor in Neisseria gonorrhoeae. Infect Inimun 63 (8): 2958-2967.

26. Legrain M, et al. 1993. Cloning and characterization of Neisseria meningitides genes encoding the transferrin-binding proteins Tbpl and Tbp2. Gene 130 (1): 73-80.

27. Cornelissen CN, Biswas GD, Sparling PF. 1993. Expression of gonococcal transferrin-binding protein 1 causes Escherichia coli to bind human transferrin. J Bacteriol 175 (8): 2448-2450.

28. Anderson JE, Sparling PF, Cornelissen CN. 1994. Gonococcal transferrin-binding protein 2 facilitates but is not essential for transferrin utilization. J Bacteriol 176 (11): 31623170.

29. Ogunnariwo JA, Schryvers AB. 1996. Rapid identification and cloning of bacterial transferrin and lactoferrin receptor protein genes. J Bacteriol 178 (24): 7326-7328.

30. Loosmore SM, et al. 1996. Cloning and expression of the Haemophilus influenzae transferrin receptor genes. Mol Microbiol 19 (3): 575-586.

31. Pettersson, A. et al. 1993. Molecular Characterization of the 98-Kilodalton Iron-Regulated Outer membrane Protein of Neisseria meningitides. Infect. Immun. 61 (ti): 47244733.

32. Ogunnariwo, J.A., Woo, T.K.W., Lo, R.Y.C., Gonzalez, G.C., and Schryvers, A.B. (1997) Characterization of the *Pasteurella haemolytica* transferrin receptor genes and the recombinant receptor proteins. *Microbial Pathog* **23**:273-284.

33. Yang, Y.P., Myers, L.E., McGuinness, U., Chong, P., Kwok, Y., Klein, M.H., and Harkness, R.E. (1997) The outer membrane protein, CD, extracted from *Moraxella (Branhamella) catarrhalis* is a potential vaccine antigen that induces

bactericidal antibodies. *FEMS Immun Med Microbiol* **17**:187-199.

34. Pettersson, A., Klarenbeek, V., van Deurzen, J., Poolman, J.T., and Tommassen, J. (1994a) Molecular characterization of the structural gene for the lactoferrin receptor of the meningococcal strain H44/76. *Microb Pathog* **17**:395-408.

35. Needleman, S.B., and Wunsch, C.D. 1970, J. Mol Biol. 48:443-453.

36. Sellers, P.H. 1974 On the theory and computation of evolutionary distances. J. Appl. Math(Siam) 26:787-793.

37. Waterman, M.S., Smith, T.F., and Beyer, W.A. 1976. Advan. Math. 20:367-387.

38. Smith, T.F., and Waterman, M.S. 1981 Identification of common molecular subsequences. J. Mol. Biol. 147:195-197.

39. Sobel, E. and Martinez, H.M. 1985 A Multiple Sequence Alignment Program. Nucleic Acid Res. 14:363-374.

40. Bonnat, R.A., Yu, R.H. and Schryvers, A.B. 1995, Biochemical Analysis of Lactoferrin Receptors in the *Neisseriaceae*: Identification of a Second Lactoferrin Receptor Protein. Microb. Pathog. 19:285-297.

CLAIMS

What we claim is:

1. A purified and isolated nucleic acid molecule encoding a lactoferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of the lactoferrin receptor protein.

2. The nucleic acid molecule of claim 1 wherein the lactoferrin receptor protein is the lactoferrin receptor binding protein 1 (Lbp1) of the *Moraxella* strain.

3. The nucleic acid molecule of claim 1 wherein the lactoferrin receptor protein is the lactoferrin receptor binding protein 2 (Lbp2) of the *Moraxella* strain.

4. The nucleic acid molecule of claim 1 wherein the lactoferrin receptor protein is the open reading frame protein 3 (ORF3) of the *Moraxella* strain.

5. The nucleic acid molecule of claim 1 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.

6. The nucleic acid molecule of claim 5 wherein the strain of *Moraxella catarrhalis* is *Moraxella catarrhalis* 4223, Q8 or VH16.

7. A purified and isolated nucleic acid molecule encoding at least one lactoferrin binding protein of *Moraxella* and having a restriction map as shown in Figure 3 for *M. catarrhalis* 4223, Figure 5 for *M. catarrhalis* Q8 or Figure 17 for *M. Catarrhalis* VH19 or the equivalent map for another strain of *M. catarrhalis*.

8. A purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:

(a) a DNA sequence as set out in Figure 2 or 4 (SEQ ID Nos. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 69) or the complementary DNA sequence thereto;

(b) a DNA sequence encoding an amino acid sequence as set out in Figure 2 or 4 (SEQ ID Nos. 11, 12, 13, 14, 15, 16, 17, 18, 70) or the complementary DNA sequence thereto; and

(c) a DNA sequence encoding a functional lactoferrin receptor protein of *Moraxella*.

9. The nucleic acid molecule of claim 8, wherein the DNA sequence defined in (c) hybridizes under stringent conditions to any one of the sequences defined in (a) or (b).

10. The nucleic acid molecule of claim 8, wherein the DNA sequence defined in (c) is that encoding the equivalent lactoferrin receptor protein from another strain of *Moraxella*.

11. A vector adapted for transformation of a host comprising the nucleic acid molecule of claim 1 or 8.

12. The vector of claim 11 encoding at least a fragment of a lactoferrin receptor protein and having the characteristics of a plasmid selected from the group consisting of pLD3, pLDW3, pLD1-8 (ATCC 97,997), pLDW1 (ATCC 97,998), pVH19pcr1 and pVH19pcr2.

13. The vector of claim 11 further comprising expression means operatively coupled to the nucleic acid molecule for expression of said lactoferrin receptor protein of a strain of *Moraxella* or the fragment or the analog of the lactoferrin receptor protein by the host containing the vector.

14. The vector of claim 13 having the characteristics of plasmid pRD1A, pRD1B, pQW1A, PQW1B, pRD2B, pQW2B, pLRD3 and pLQW3.

15. A transformed host containing an expression vector as claimed in claimed 13.

16. A method of forming a substantially pure recombinant lactoferrin receptor protein, which comprises:

growing the transformed host of claim 15 to express a lactoferrin receptor protein as inclusion bodies,

purifying the inclusion bodies free from cellular material and soluble proteins,

solubilizing lactoferrin receptor protein from the purified inclusion bodies, and

purifying the lactoferrin receptor protein free from other solubilized materials.

17. The method of claim 16 wherein said lactoferrin receptor protein comprises Lbp1 alone, Lbp2 alone, ORF3 alone or a mixture of two or more of Lbp1, Lbp2 and ORF3.

18. The method of claim 17 wherein said lactoferrin receptor protein is a least about 70% pure.

19. The method of claim 18 wherein said lactoferrin receptor protein is a least about 90% pure.

20. A recombinant lactoferrin receptor protein or fragment or analog thereof producible by the transformed host of claim 14.

21. The protein of claim 20 which is lactoferrin receptor binding protein 1 (Lbp1) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

22. The protein of claim 20 which is lactoferrin receptor binding protein 2 (Lbp2) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

23. The protein of claim 20 which is open reading frame protein 3 (ORF3) of the *Moraxella* strain devoid of other proteins of the *Moraxella* strain.

24. The protein of claim 20 wherein the strain of *Moraxella* is a strain of *Moraxella catarrhalis*.

25. An open reading frame protein 3 (ORF3) of a *Moraxella* strain or a fragment or analog of the lactoferrin binding protein.

26. The protein of claim 25 wherein said *Moraxella* strain is a strain of *M. catarrhalis*.

27. The protein of claim 26 wherein said *M. catarrhalis* strain is *M. catarrhalis* 4223 or Q8.

28. The protein of claim 24 having a deduced amino acid as set forth in Figure 2 or 4 (SEQ ID No: 14, 18).

29. An immunogenic composition, comprising at least one active component selected from the group consisting of:

(A) a purified and isolated nucleic acid molecule encoding a lactoferrin receptor protein of a strain of *Moraxella* or a fragment or an analog of lactoferrin receptor protein;

(B) a purified and isolated nucleic acid molecule having a DNA sequence selected from the group consisting of:

(a) a DNA sequence as set out in Figure 2 or 4 (SEQ ID Nos. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 69) or the complementary DNA sequence thereto;

(b) a DNA sequence encoding an amino acid sequence as set out in Figure 2 or 4 (SEQ ID Nos. 11, 12, 13, 14, 15, 16, 17, 18, 70) or the complementary DNA sequence thereto; and

(c) a DNA sequence encoding a functional lactoferrin receptor protein of *Moraxella*;

(C) a recombinant lactoferrin receptor protein or fragment or analog thereof producible by a transformed host containing an expression vector comprising a nucleic acid molecule as defined in (A) or (B) and expression means operatively coupled to the nucleic acid molecule for expression by the host of the recombinant lactoferrin receptor protein or fragment or analog thereof; or

(D) an open reading frame protein 3 (ORF3) of a *Moraxella* strain or a fragment or analog of the open reading frame protein;

and a pharmaceutically acceptable carrier therefor, said at least one active component producing an immune response when administered to a host.

30. A method for generating an immune response in a host, comprising administering to the host an immunoeffective amount of the immunogenic composition of claim 29.

31. A method of determining the presence, in a sample, of nucleic acid encoding a lactoferrin receptor protein of a strain of *Moraxella* , comprising the steps of:

(a) contacting the sample with the nucleic acid of claim 1 or 7 to produce duplexes comprising the nucleic acid molecule and any nucleic acid molecule encoding the lactoferrin receptor protein of a strain of Moraxella present in the sample and specifically hybridizable therewith; and

(b) determining production of the duplexes.

32. A diagnostic kit for determining the presence, in a sample of nucleic acid encoding a lactoferrin receptor protein of a strain of *Moraxella*, comprising:

(a) the nucleic acid molecule of claim 1 or 7;

(b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the

nucleic acid molecule and any said nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

(c) means for determining production of the duplexes.

FIG.1A

Alignment of translated 2.2kb *lbpA* PCR fragments

```
MNQSKQNNKSKKSKQVLKLSALSLGLLNITQVALANTTADK                  Tbp1
MSKSITKTQTPSVHTMTTHRLNLAIKAALFGVAVLPLSVWAQENTQTDAN        Lbp1

AEATDKTNLVVVLDETVVTAKKNAPVSRKANEVTGLGKVVKTAETINKEQ        Tbp1
SDAKDTKTPVVYLDAITVTAAPSA---RFDTDVTGLGKTVKTADTLAKEQ        Lbp1

VLNIRDLTRYDPGIAVVEQGRGASSGYSIRGMDKNRVAVLVDGINQAQHY        Tbp1
VQGIRDLVRYETGVSVVEQGRGGSSGFAIHGVDKNRVGITVDGIAQIQSY        Lbp1

--QGPVAGKNYAAGGAINEIEYENVRSVEISKGANSSEYGSGALSGSVAFVT      Tbp1
ALKDESTKRAGAGSGAMNEIEIENIAAVAINKGGNALEAGSGALGGSVAFHT      Lbp1

KTADDIIKDGKDWGVQTKTAYASKNNAWVNSVAAAGKAGSFSGLIIYTDR       Tbp1
KDVSDVLKSGKNLGAQSKTTYNSKNDHFSQTLAAAGKTERVEAMVQYTYR       Lbp1
                                                 QYT-R     PCR4
                                                 QYT-R     PCR5

RGQEYKAHDDAYQGSQSFDRAVATTDPNNRTFLIANECANGNYEACAAGG        Tbp1
KGKENKAHSDLNGINQSLYRLGAWQQKYDLRKPNELFAGTSYITESCLAS        Lbp1
KG-ENKAHSDLNGINQSLYRLGAWQQKYDLRKPNELFAGTSYITESCLAS        PCR4
KG-ENKAHSDLNGINQSLYRLGAWQQKYDLRKPNELFAGTSYITESCLAS        PCR5

QTKLQAKPTNVRDKVNVKDYTGPNRLIPNPLTQDSKSLLLRPGYQLNDKH        Tbp1
DDPKSCVQYPYVYTKARPDGIGNRNFSELSDAEKAQYLASTHPHEVVSAK        Lbp1
DDPKSCVQYPYVYTKARPDGIGNRNFSELSDAEKAQYLASTHPHEVVSAK        PCR4
DDPKSCVQYPYVYTKARPDGIGNRNFSELSDAEKAQYLASTHPHEVVSAK        PCR5
```

FIG.1B

```
YVGGVYEITKQNYAMQDKTVPAYLAVHDIEKSRLSNHAQANGYYQGNNLGERIRDT     Tbp1
DYTGIYRLLPDPMDYRSDSYLARLNIKITPNLVSKLLLEDTKQTYNIRDM     Lbp1
DYTGIYRLLPDPMDYRSDSYLARLNIKITPNLVXKLLLEDTKQTYNIRDM     PCR4
DYTGTYRLLPDPMDYRSDSYLARLNIKITPNLVSKLLLEDTKQTYNIRDM     PCR5

IGPDSGYGINYAHGVFYDEKHQKDRLGLEYVYDSKGENKWFDDVRVSYDKQDIT       Tbp1
RHCSYHGARLGNDGKPANGGSIVLCDDYQEYLNANDASQALFRPGANDAP     Lbp1
RHCSYHGARLGNDGKPANGGSIVLCDDYQEYLNANDASQALFRPGANDAP     PCR4
RHCSYHGARLGNDGKPANGGSIVLCDDYQEYLNANDASQASFRPGANDAP     PCR5

LRSQLTNTHCSTYPHIDKNCTPDVNKPFSVKEVDNNAYKEQHNLIKAVFN     Tbp1
IPKLAYARSSVFNQEHGKTRYGLSFEFKPDTPWFKQAKLNLHQQNIQIIN     Lbp1
IPKLAYARSSVFNQEHGKTRYGLSFEFKPDTPWFKQAKLNLHQQNIQIIN     PCR4
IPKLAYARSSVFNQEHGKTRYGLGFEFKPDTPWFKQAKLNLHQQNIQIIN     PCR5

KKMALGSTHHHINLQVGYDKFNSSLSRVEYRLATHQSYQKLDYTPPSNPL     Tbp1
HDIKKSCSQYPKVDLNCGISEIGHYEYQNNYRYKEGRASLTGKLDFNFDL     Lbp1
HDIKKSCSQYPKVDSNCGISEIGHYEYQXNYRYKEGRASLTGKLDFNFDL     PCR4


PDKFKPILGSNNKPICLDAYGYGHDHPQACNAKNSTYQNFAIKKGIEQYN     Tbp1
LGQHDLTVLAGADKVKSQFRANNPRRTIIDTTQGDAIIDESTLTAQEQAK     Lbp1
LGQHDLTVLAGTDKVKSQFRANNPRRTIIDTTQGDAIIDESTLTAQEQAK     PCR4
```

FIG.1C

```
QKTNTDKIDYQAIIDQYDKQNPNSTLKPFEKIKQSLGQEKYNKIDELGFK  Tbp1
FKQSGAAWIVKNRLGRLEEKDACGNANECERAPIHGSNQYVGINNLYTPN  Lbp1
FKQSGAAWIVKNRLGRLEEKDACGNANECERAPIHGSNQYVGINNLYTPN  PCR4

AYKDLRNEWAGWTNDNSQQNANKGTDNIYQPNQATVVKDDKCKYSETNSY  Tbp1
DYVDLSFGGRLDKQRIHSTDSNIISKTYTNKSYNFGAAVHLTPDFSLLYK  Lbp1
DYVDXSFGGRLDKQRIHSTDSNIISKTYTNKSYNFGAAVHLTPDFSLLYK  PCR4
          TDSNIISKTYTNKSYNFGAAVHXTPDFSLLYK                PCR5

ADCSTTRHISGDNYFIALKDNMTINKYVDLGLGARYDRIKHKSDVPLVDNSASNQLSWNFGVV  Tbp1

TAKGFRTPSFYELYNYNSTAAQHKNDPDVSFPKRAVDVKPETSNTNEYGF  Lbp1
TAKGFRTPSFYELYNYNSTAAQHKNDPDVSFPKRAVDVKPETSNTNEYGF  PCR4
TAKGFRTPSFYELYNYNSTAAQHKNDPDVSFPKRAVDVKPETSNTNEYGF  PCR5

VKPTNWLDIAYRSSQGFRMPSFSEMYGERFGVTIGKGTQHGCKGLYYICQQTV  Tbp1
RYQHPWGDVEMSMFKSRYKDMLDKAIPNLTKAQQEYCKAHLDSNECVGNP  Lbp1
RYQHPWGDVEMSMFKSRYKDMLDKAIPNLTKAQQEYCRAHLDSNECVGNP  PCR4
RYQHPWGDIEMSMFKSRYKDMLDKAIPNLTKAQQEYCKAHLDSNECVGNP  PCR5

HQTKLKPEKSFNQEIGATLHNHLGSLEVSYFKNRYTDLIVGKSEEIRTLT  Tbp1
PTPKTSDEVFANLYNATIKGVSVKGKLDLHAMTSKLPDGLEMTLGYGHTK  Lbp1
PTPKTSDEVFANLYNATIKGVSVKGKLDLHAMTSKLPDGLEMTLGYGHTK  PCR4
PTPKTSDEVFANLYNATIKGVSVKGKLDLHAMTSKLPDGLEMTLGYGHTK  PCR5

QGDNAGKQRGKGDLGFHNGQDADLTGINILGRLDLNAANSRLPYGLYSTL  Tbp1
```

FIG.1D

```
AYNKVDVKGKTLNPTLAGTNILFDAIQPSRYVVGLGYDAPSQKWGANAIF   Tbp1
LGKFDYIAPKDADGWYQARPAFWDAITPARYVVGLNYDHPSQVWGIGTTL   Lbp1
LGKFXYIAPKDADGWYQARPAFWDAITPARYVVGLNYDHPSQVWGIGATL   PCR4
LGKFXYIAPKDADGWYQARPAFWDAITPARYVVGLNYDHPSQVWGIGTTL   PCR5

HSDAKNPSELLADKNLGNGNIQTKQATKAKSTPWQTL-DLSGYVNIKDNFT   Tbp1
THSKQKDENELSALRIR-NGKRETQTLTHTIPKAYTLLDMTGYYSPTESIT   Lbp1
THSKQKDENELSALRIR-NGKRETQTLTHTIPKAYTLLDMTGYYSPTESIT   PCR4
THSKQKDENELSALRIR-NGKREIQTLTHTIPKAYTLLDMTGYYSPTESIT   PCR5

LRAGVYNVFNTYYTTWEALRQTAKGAVNQHTGLSQDKHYGRYAAPGRNYQLALEMKF*   Tbp1
ARLGINNVLNTRYTTWEAARQ-------LPSEAASSTQSTRYIAPGRSYFASLEMKF*   Lbp1
ARLGINNVLNTRYTTWEAARQ-------LPSEAASSTQSTRYIAPGRSYFASLEMKF*   PCR4
ARLGINNVLNTRYTTWEAARQ-------LPSEAASSTQSTRYIAPGRSYFASLEMKF*   PCR5
```

FIG.2A

M. catarrhalis 4223 *lfr sequence*

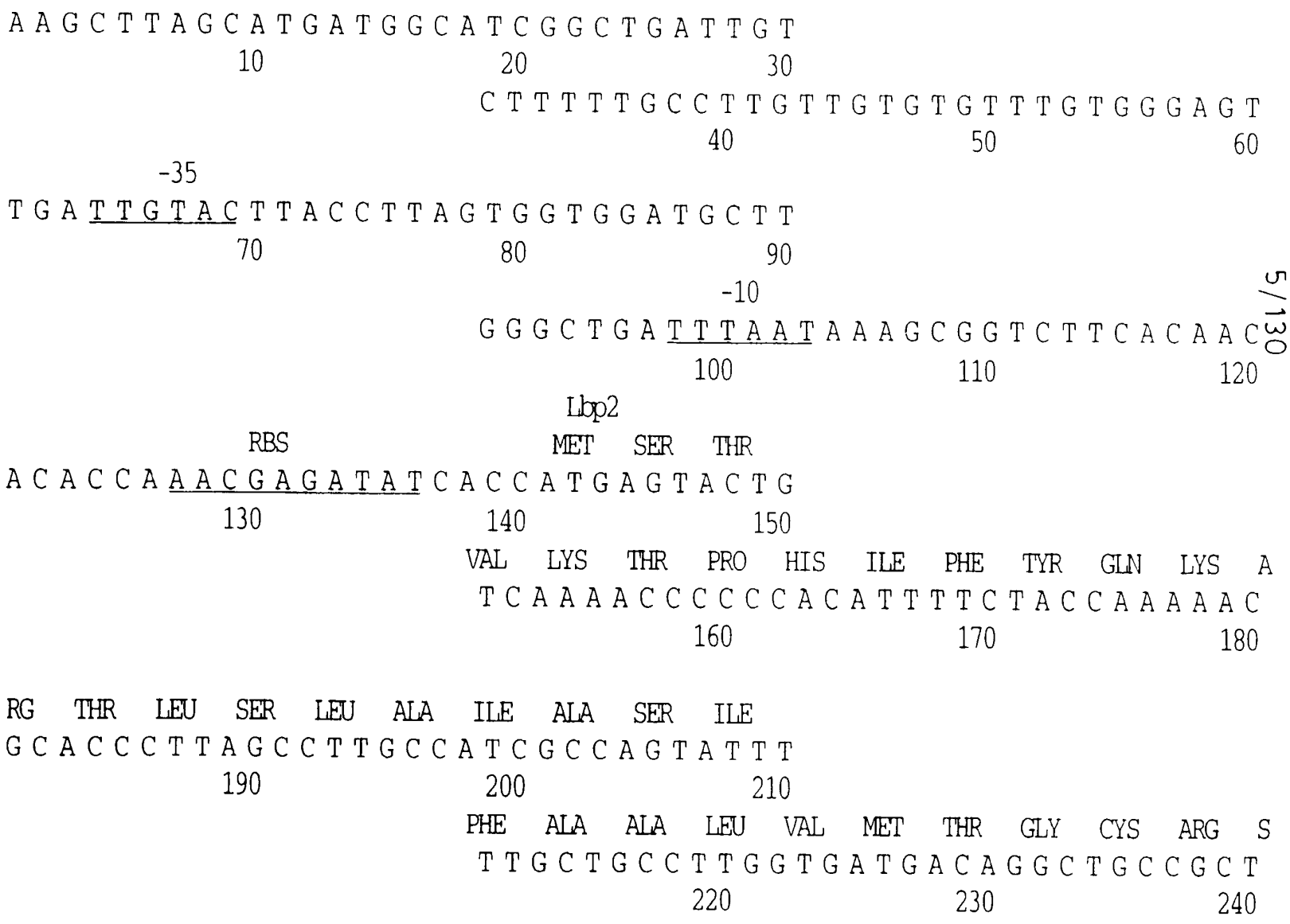


FIG.2B

```
ER   ASP   ASP   ILE   SER   VAL   ASN   ALA   PRO   ASN
C T G A T G A C A T C A G C G T C A A T G C A C C C A A T G
               250                 260                 270
                                  VAL   THR   GLN   LEU   PRO   GLN   GLY   THR   VAL   SER   P
                                   T T A C C C A A C T G C C C C A A G G C A C G G T T T C A C
                                                      280                 290                 300

RO   ILE   PRO   ASN   THR   GLY   HIS   ASP   ASN   THR
C A A T A C C G A A C A C A G G T C A T G A C A A C A C C A
               310                 320                 330
                                  ASN   ASN   THR   ASN   ASN   GLN   GLY   ASN   ASN   THR   A
                                   A T A A C A C C A A C A A T C A G G G C A A C A A C A C G G
                                                      340                 350                 360

SP   ASN   SER   THR   SER   THR   THR   ASP   PRO   ASN
A T A A C A G C A C C A G C A C A A C T G A C C C A A A T G
               370                 380                 390
                                  GLY   ASP   ASN   ASN   GLN   LEU   THR   GLN   ALA   GLN   L
                                   G C G A T A A C A A C C A A C T G A C A C A A G C A C A A A
                                                      400                 410                 420
```

FIG.2C

```
YS  THR ALA ALA ALA ALA GLY PHE PHE VAL
A G A C C G C C G C C T G C C C A G G G T T T T T T G T G A
          430       440       450

MET GLY LYS ILE ARG ASP THR SER PRO LYS A
T G G G T A A A A T T C G T G A T A C C A G C C C A A A A A
      460       470       480

SN  ASP PRO ASP TYR SER ASN ASP LEU VAL
A T G A C C C C A G A T T A T A G C A A T G A T T T A G T A C
          490       500       510

GLN GLN TRP GLN GLY LYS LEU TYR VAL GLY I
A G C A G T G G G C A A G G C A A A T T A T G T T G G T A
      520       530       540

LE  ASP ALA HIS ARG PRO ASP GLY ILE GLY
T T G A T G C C C C A T C G C C C C A G A T G G G C A T C G G G C A
          550       560       570

THR GLY LYS ASN LEU ARG GLN PRO ILE THR A
C A G G G T A A A A A C T T G C G T C A G G C C C A T C A C C G
      580       590       600
```

FIG.2D

```
LA  ASN  ASP  ILE  LYS  PRO  LEU  TYR  PHE  ASN
C C A A T G A C A T C A A A A C C C C T T G T A T T T T A A C A
                610                  620                  630

                LYS  PHE  PRO  ALA  LEU  SER  ASP  LEU  HIS  LEU  A
                A T T C C C C T G C A T T G T C T G A T T T G C A T T A G
                          640                  650                  660

SP  SER  GLU  ARG  HIS  ARG  PHE  ASP  PRO  LYS
A C A G T G A A C G C C A C C G T T T T G A C C C C C A A A A
          670                  680                  690

                LYS  LEU  ASN  THR  ILE  LYS  VAL  TYR  GLY  TYR  G
                A G C T A A A C A C C A T T A A A G T G T A T G G T T A T G
                          700                  710                  720

LY  ASN  LEU  THR  THR  PRO  SER  LYS  ASN  ASN
G C A A C T T A A C A A C A C C C C T C T A A A A A C A A C A
          730                  740                  750

                THR  TYR  ILE  ASN  HIS  GLN  GLN  ALA  ASP  ASN  L
                C T T A C A T C A A T C A T C A G C A A G C T G A T A A T A
                          760                  770                  780
```

FIG.2E

```
YS  LYS  ASN  ASN  LYS  PRO  VAL  ASP  PRO  TYR
 A G A A A A A T A A C A A G C C T G T T G A C C C T T A T
              790                  800                  810
                                  GLU  ASN  ILE  ARG  PHE  GLY  TYR  LEU  GLU  LEU  G
                                   G A A A A T A T C C G T T T T G G G T A T C T T G A A C T A C
                                                       820                  830                  840

LN  GLY  SER  SER  LEU  THR  GLN  LYS  ASN  ALA
A A G G A A G C A G T C T G A C C C A A A A A A A T G C C G
              850                  860                  870
                                  ASP  THR  PRO  ASN  ASP  LYS  ASP  ARG  ILE  PRO  L
                                   A T A C T C C A A A T G A C A A A G A C C G C A T T C C C A
                                                       880                  890                  900

YS  PRO  MET  PRO  ILE  LEU  PHE  TYR  HIS  GLY
 A A C C C A T G C C C A T T T T G T T T T A T C A C G G A G
              910                  920                  930
                                  GLU  ASN  ALA  SER  SER  GLN  LEU  PRO  SER  ALA  G
                                   A A A A C G C C A G C A G C C A G C T G C C C A G T G C T G
                                                       940                  950                  960
```

FIG.2F

```
LY  LYS  PHE  ASN  TYR  THR  GLY  ASN  TRP  LEU
G T A A A T T T A A C T A C A C A G G C A A C T G G C T G T
          970                 980                 990

                    TYR  LEU  SER  ASP  VAL  LYS  LYS  ARG  PRO  ALA  L
                    A C C T A A G T G A T G T C A A A A A A C G C C C T G C A C
                              1000                1010                1020

EU  SER  ALA  SER  ASP  ASP  ARG  VAL  GLY  VAL
T T T C A G C A T C A G A T G A T C G A G T G G G G G T C T
          1030                1040                1050

                    TYR  LEU  ASN  ALA  SER  GLY  LYS  SER  ASN  GLU  G
                    A T C T C A A T G C C C A G T G G C A A A T C C A A T G A G G
                              1060                1070                1080

LY  ASP  VAL  SER  ALA  ALA  HIS  ILE  TYR
G C G A T G T C G T C A G T G C C C C A C A T T T A T C
          1090                1100                1110

                    LEU  ASN  GLY  PHE  GLN  TYR  LYS  HIS  THR  PRO  A
                    T A A A C G G G C T T T C A A T A T A A G C A C A C G C C T G
                              1120                1130                1140
```

FIG.2G

```
LA  THR  TYR  GLN  VAL  ASP  PHE  ASP  THR  ASN
C C A C T T A T C A G G T G G A T T T T G A C A C A A A C T
          1150            1160            1170
                          SER  LEU  THR  GLY  LYS  LEU  SER  TYR  TYR  ASP  A
                           C A T T A A C A G G C A A G C T G T C T T A T T A T G A C A
                                      1180            1190            1200

SN  PRO  ASN  GLN  GLN  THR  ALA  GLN  GLY  LYS
A T C C C A A C C A G C A A A C T G C C C A A G G C A A A T
          1210            1220            1230
                          TYR  ILE  LYS  SER  GLN  PHE  ASP  THR  THR  LYS  L
                           A C A T C A A A A G C C A A T T T G A C A C T A C C A A A A
                                      1240            1250            1260

YS  VAL  ASN  GLU  THR  ASP  VAL  TYR  GLN  ILE
A A G T C A A T G A A A C C G A T G T G T A T C A A A T T G
          1270            1280            1290
                          ASP  ALA  LYS  ILE  ASN  GLY  ASN  ARG  PHE  VAL  G
                           A T G C C A A A A T C A A C G G C A A C C G C T T C G T C G
                                      1300            1310            1320
```

FIG.2H

```
LY   THR   ALA   LYS   SER   LEU   VAL   ASN   GLU   ASN
G T A C G G C C A A A T C T T T G G T T A A T G A G A A C A
            1330                1340                1350
                                THR   GLU   THR   ALA   PRO   PHE   ILE   LYS   GLU   LEU   P
                                 C A G A A A C C G C A C C T T T T A T C A A A G A G C T G T
                                              1360                1370                1380

HE   SER   LYS   LYS   ALA   ASN   PRO   ASN   ASN   PRO
T C T C C A A A A A A G C C A A T C C C A A T A A C C C A A
            1390                1400                1410
                                ASN   PRO   ASN   SER   ASP   THR   LEU   GLU   GLY   GLY   P
                                 A C C C T A A T T C A G A C A C G C T A G A A G G C G G G T
                                              1420                1430                1440

HE   TYR   GLY   GLU   SER   GLY   ASP   GLU   LEU   ALA
T T T A T G G T G A G T C G G G C G A T G A G C T G G C G G
            1450                1460                1470
                                GLY   LYS   PHE   LEU   SER   ASN   ASP   ASN   ALA   SER   T
                                 G T A A A T T T T T A T C C A A T G A C A A C G C A T C T T
                                              1480                1490                1500
```

FIG.2I.

```
YR  VAL  VAL  PHE  GLY  GLY  LYS  ARG  ASP  LYS
A T G T G G T C T T T G G T G G T A A A C G A G A C A A A A
            1510              1520              1530
                          THR  ASP  LYS  PRO  VAL  ALA  THR  LYS  THR  VAL  T
                           C A G A C A A A C C T G T C G C C A C A A A A A C G G T G T
                                        1540              1550              1560

YR  PHE  SER  ALA  GLY  PHE  GLU  LYS  PRO  SER
A T T T T A G T G C A G G C T T T G A A A A A C C T A G C A
            1570              1580              1590
                          THR  SER  PHE  VAL  ASP  ASN  GLU  THR  ILE  GLY  A
                           C C A G T T T T G T G G A T A A T G A A A C G A T T G G C A
                                        1600              1610              1620

RG  ILE  ILE  ASN  SER  LYS  LYS  LEU  ASN  ASP
G A A T T A T T A A C A G C A A A A A G T T A A A T G A T G
            1630              1640              1650
                          ALA  VAL  ASN  GLU  LYS  ILE  ASP  ASN  GLY  ASP  I
                           C G G T G A A T G A G A A A A T T G A T A A T G G T G A T A
                                        1660              1670              1680
```

FIG.2J

```
LE   PRO   THR   SER   ASP   GLU   ARG   TYR   ASP   GLU
T T C C T A C C A G T G A T G A A C G C T A T G A T G A A T
               1690                1700                1710
                                  PHE   PRO   TRP   GLY   GLU   LYS   LYS   ALA   GLU   PHE   T
                                  T T C C T T G G G G C G A A A A A A A A G C A G A A T T C A
                                                      1720                1730                1740

HR   LYS   LYS   VAL   SER   SER   SER   THR   GLN   ALA
C C A A A A A A G T C A G C A G C A G C A C C C A A G C C G
               1750                1760                1770
                                  VAL   PRO   ALA   TYR   PHE   GLY   GLN   HIS   ASP   LYS   P
                                  T G C C A G C T T A T T T T G G G C A A C A T G A T A A A T
                                                      1780                1790                1800

HE   TYR   PHE   ASN   GLY   ASN   TYR   TYR   ASP   LEU
T T T A T T T T A A T G G C A A C T A T T A T G A C C T A T
               1810                1820                1830
                                  SER   ALA   SER   SER   VAL   ASP   LYS   LEU   ALA   PRO   A
                                  C A G C C A G C A G T G T T G A T A A A T T G G C C C C T G
                                                      1840                1850                1860
```

FIG.2K

```
LA  ASP  ALA  VAL  LYS  ALA  ASN  GLN  SER  ILE
C C G A T G C T G T C A A A G C C A A C C A A T C C A T T A
              1870                 1880                 1890
                                  LYS  GLU  LYS  TYR  PRO  ASN  ALA  THR  LEU  ASN  L
                                   A A G A A A A A T A C C C T A A T G C C A C A C T A A A T A
                                                      1900                 1910                 1920

YS  ASP  ASN  GLN  VAL  THR  ALA  ILE  VAL  LEU
A G G A C A A C C A A G T T A C C G C C A T C G T G C T A C
              1930                 1940                 1950
                                  GLN  GLU  ALA  LYS  ASP  ASN  LYS  PRO  TYR  THR  A
                                   A A G A A G C C A A A G A T A A T A A G C C T T A T A C C G
                                                      1960                 1970                 1980

LA  ILE  ARG  ALA  LYS  SER  TYR  GLN  HIS  ILE
C C A T T C G T G C C A A A A G C T A T C A G C A C A T C A
              1990                 2000                 2010
                                  SER  PHE  GLY  GLU  THR  LEU  TYR  ASN  ASP  ALA  A
                                   G T T T TG G C G A G A C G C T G T A T A C G A T G C C A A
                                                      2020                 2030                 2040
```

FIG.2L

SN GLN THR PRO THR ARG SER TYR PHE VAL
A C C A A A C C C C A A C A C G C A G T T A T T T T G T G C
2050 2060 2070

GLN GLY GLY ARG ALA ASP THR SER THR THR L
A A G G C G G T A G G G C A G A T A C C A G C A C C A C G C
2080 2090 2100

EU PRO LYS ALA GLY LYS PHE THR TYR ASN
T G C C C A A G G C A G G T A A A T T C A C T T A C A A C G
2110 2120 2130

GLY LEU TRP ALA GLY TYR LEU ILE GLN LYS L
G T C T T T G G G C A G G C T A T C T T A T C C A A A A A A
2140 2150 2160

YS ASP LYS GLY TYR SER ASN ASN GLU GLU
A G G A C A A A G G T T A T A G C A A T A A T G A A G A A A
2170 2180 2190

THR ILE LYS LYS LYS GLY HIS GLN ASP TYR L
C C A T C A A G A A A A A A G G C C A T C A A G A T T A T C
2200 2210 2220

FIG.2M

```
EU   LEU   THR   GLU   ASP   PHE   THR   PRO   GLU   ASP
T G T T A A C C G A A G A C T T C A C C C C A G A A G A T G
             2230                2240                2250
                                ASP   ASP   ASP   ASP   LEU   THR   ALA   SER   ASP   ASP   S
                                 A T G A C G A T G A T T T G A C C G C A T C T G A T G A T T
                                                   2260                2270                2280

ER   GLN   ASP   ASP   ASP   ALA   HIS   GLY   ASP   ASP
C A C A A G A T G A T G A T G C A C A T G G C G A T G A T G
             2290                2300                2310
                                ASP   LEU   ILE   ALA   SER   ASP   ASP   SER   GLN   ASP   A
                                 A T T T G A T T G C A T C T G A T G A T T C A C A A G A T G
                                                   2320                2330                2340

SP   ASP   ALA   ASP   GLY   ASP   ASP   ASP   SER   ASP
A T G A C G C A G A T G G C G A T G A C G A T T C A G A T G
             2350                2360                2370
                                ASP   LEU   GLY   ASP   GLY   ALA   ASP   ASP   ALA   ALA   A
                                 A T T T G G G T G A T G G T G C A G A T G A C G C C G C C G
                                                   2380                2390                2400
```

FIG.2N

```
LA  GLY  LYS  VAL  TYR  HIS  ALA  GLY  ASN  ILE
C A G G C A A A G T G T A T C A T G C A G G T A A T A T T C
             2410                 2420                 2430
                                 ARG  PRO  GLU  PHE  GLU  ASN  LYS  TYR  LEU  PRO  I
                                  G C C C T G A A T T T G A A A A C A A A T A C T T G C C C A
                                                2440                 2450                 2460

LE  ASN  GLU  PRO  THR  HIS  GLU  LYS  THR  PHE
T T A A T G A G C C T A C T C A T G A A A A A A C C T T T G
             2470                 2480                 2490
                                 ALA  LEU  ASP  GLY  LYS  ASN  LYS  ALA  LYS  PHE  A
                                  C C C T A G A T G G T A A A A A T A A A G C T A A G T T T G
                                                2500                 2510                 2520

SP  VAL  ASP  PHE  ASP  THR  ASN  SER  LEU  THR
A T G T G G A T T T T G A C A C C A A C A G C C T A A C T G
             2530                 2540                 2550
                                 GLY  LYS  LEU  ASN  ASP  GLU  ARG  GLY  ASP  ILE  V
                                  G T A A A T T A A A C G A T G A G A G A G G T G A T A T C G
                                                2560                 2570                 2580
```

FIG.2O

```
AL   PHE  ASP  ILE  LYS  ASN  GLY  LYS  ILE  ASP
T C T T T G A T A T C A A A A A T G G C A A A A T T G A T G
              2590              2600              2610
                               GLY  THR  GLY  PHE  THR  ALA  LYS  ALA  ASP  VAL  P
                                G C A C A G G C T T T A C C G C C A A A G C C G A T G T G C
                                                  2620              2630              2640

RO   ASN  TYR  ARG  GLU  GLU  VAL  GLY  ASN  ASN
C A A A C T A T C G T G A A G A A G T G G G T A A C A A C C
              2650              2660              2670
                               GLN  GLY  GLY  GLY  PHE  LEU  TYR  ASN  ILE  LYS  A
                                A A G G T G G C G G T T T C T T A T A C A A C A T C A A A G
                                                  2680              2690              2700

SP   ILE  ASP  VAL  LYS  GLY  GLN  PHE  PHE  GLY
A T A T T G A T G T C A A G G G G C A A T T T T T T G G C A
              2710              2720              2730
                               THR  ASN  GLY  GLU  GLU  LEU  ALA  GLY  GLN  LEU  G
                                C A A A T G G C G A A G A G T T G G C A G G G C A G T T A C
                                                  2740              2750              2760
```

FIG.2P

```
LN   TYR   ASP   LYS   GLY   ASP   GLY   ILE   ASN   ASP
AGTACGACAAAGGCGATGGCATCAATGACA
           2770          2780          2790
                       THR   ALA   GLU   LYS   ALA   GLY   ALA   VAL   PHE   GLY   A
                       CCGCCGAAAAAGCAGGGGCTGTCTTTGGGG
                                   2800          2810          2820

LA   VAL   LYS   ASP   LYS   ***
CTGTTAAAGATAAATAAAGCCCCCTTCATC
           2830          2840          2850
                       ATCGTTTAGTCGCTTGACCGACAGTTGATG
                                   2860          2870          2880

ACGCCCTTGGCAATGTCTTAAAACAGCACT
           2890          2900          2910
                       TTGAAACAGTGCCTTGGGCGAATTCTTGGA
                                   2920          2930          2940

TAAATGCACCAGATTTGCCTTGGGCTAATA
           2950          2960          2970
                             -35                   -10
                       TCTTGATAAAACATCGCCATAAAAATAGAAA
                                   2980          2990          3000
```

FIG.2Q

```
                                        Lbp1
                RBS                     MET  SER  LYS
A T A A A G T T T A G G A T T T T T T T A T G T C A A A A T
          3010              3020              3030
                          SER  ILE  THR  LYS  THR  GLN  THR  PRO  SER  VAL  H
                           C T A T C A C A A A A A C A C A A A C A C C A T C A G T C C
                                     3040              3050              3060

          2nd possible start
IS  THR  MET  THR  THR  HIS  ARG  LEU  ASN  LEU
A T A C C A T G A C C A C G C A C C G C T T A A A C C T T G
          3070              3080              3090
                          ALA  ILE  LYS  ALA  ALA  LEU  PHE  GLY  VAL  ALA  V
                           C C A T C A A A G C G G C G T T A T T T G G T G T G G C A G
                                     3100              3110              3120

AL  LEU  PRO  LEU  SER  VAL  TRP  ALA  GLN  GLU
T T T T A C C C C T A T C C G T C T G G G C G C A A G A G A
          3130              3140              3150
                          ASN  THR  GLN  THR  ASP  ALA  ASN  SER  ASP  ALA  L
                           A C A C T C A G A C A G A T G C C A A C T C T G A T G C C A
                                     3160              3170              3180
```

FIG.2R

```
YS  ASP  THR  LYS  THR  PRO  VAL  VAL  TYR  LEU
A A G A C A C A A A A C C C C C T G T C G T C T A T T T A G
                3190                 3200                 3210

                ASP  ALA  ILE  THR  VAL  THR  ALA  ALA  PRO  SER  A
                A T G C C A T C A C G G T A A C C G C C C C C A T C T G
                                3220                 3230                 3240

LA  PRO  VAL  SER  ARG  PHE  ASP  THR  ASP  VAL
C C C C T G T T T C T C G G G T T T G A C A C C G A T G T A A
                3250                 3260                 3270

                THR  GLY  LEU  GLY  LYS  THR  VAL  LYS  THR  ALA  A
                C A G G G C T T G G C A A A A C G G T C A A A A C C G C T G
                                3280                 3290                 3300

SP  THR  LEU  ALA  LYS  GLU  GLN  VAL  GLN  GLY
A C A C G C T G G G C A A A A G A A C A A G T G C A G G G G C
                3310                 3320                 3330

                ILE  ARG  ASP  LEU  VAL  ARG  TYR  GLU  THR  GLY  V
                A T T C G T G A T T T G G T G C G T T A T G A A A C T G G G G
                                3340                 3350                 3360
```

FIG.2S

```
AL  SER  VAL  VAL  GLU  GLN  GLY  ARG  GLY  GLY
T G A G T G T G G T T G A G C A G G G G C G T G G T G G C A
            3370              3380              3390
                              SER  SER  GLY  PHE  ALA  ILE  HIS  GLY  VAL  ASP  L
                               G C A G C G G A T T T G C C A T T C A T G G C G T G G A T A
                                                 3400              3410              3420

YS  ASN  ARG  VAL  GLY  ILE  THR  VAL  ASP  GLY
A A A A C C G A G T G G G C A T T A C C G T A G A T G G C A
            3430              3440              3450
                              ILE  ALA  GLN  ILE  GLN  SER  TYR  LYS  ASP  GLU  S
                               T T G C C C A A A T T C A A T C C T A C A A A G A T G A A T
                                                 3460              3470              3480

ER  THR  LYS  ARG  ALA  GLY  ALA  GLY  SER  GLY
C C A C C A A A C G A G C T G G T G C A G G C T C T G G G G
            3490              3500              3510
                              ALA  MET  ASN  GLU  ILE  GLU  ILE  GLU  ASN  ILE  A
                               C G A T G A A T G A G A T A G A G A T T G A A A A C A T T G
                                                 3520              3530              3540
```

FIG.2T

```
LA  ALA  VAL  ALA  ILE  ASN  LYS  GLY  GLY  ASN
C C G C C G T T G C C A T C A A T A A A G G T G G T A A T G
           3550           3560           3570
                          ALA  LEU  GLU  ALA  GLY  SER  GLY  ALA  LEU  GLY  G
                          C C C T A G A A G C A G G C T C T G G T G C G T T G G G C G
                                    3580           3590           3600

 LY  SER  VAL  ALA  PHE  HIS  THR  LYS  ASP  VAL
G T T C G G T G G C G T T T C A T A C C A A A G A T G T G A
           3610           3620           3630
                          SER  ASP  VAL  LEU  LYS  SER  GLY  LYS  ASN  LEU  G
                          G C G A T G T C T T A A A A T C T G G T A A A A A T C T T G
                                    3640           3650           3660

LY  ALA  GLN  SER  LYS  THR  THR  TYR  ASN  SER
G C G C T C A A A G C A A A A C C A C T T A T A A C A G C A
           3670           3680           3690
                          LYS  ASN  ASP  HIS  PHE  SER  GLN  THR  LEU  ALA  A
                          A A A A T G A C C A T T T T A G T C A G A C G C T G G C A G
                                    3700           3710           3720
```

FIG.2U

```
LA  ALA  GLY  LYS  THR  GLU  ARG  VAL  GLU  ALA
C G G C A G G T A A A A C C G A G C G T G T G G A A G C G A
          3730                3740                3750

MET  VAL  GLN  TYR  THR  TYR  ARG  LYS  GLY  LYS  G
T G G T G C A A T A T A C C T A C C G T A A A G G C A A A G
          3760                3770                3780

LU  ASN  LYS  ALA  HIS  SER  ASP  LEU  ASN  GLY
A A A A C A A A G C A C A C A G C G A C C T A A A T G G C A
          3790                3800                3810

ILE  ASN  GLN  SER  LEU  TYR  ARG  LEU  GLY  ALA  T
T C A A C C A A A G C C T A T A T C G C T T G G G T G C A T
          3820                3830                3840

RP  GLN  LYS  TYR  ASP  LEU  ARG  LYS  PRO
G G C A A C A A A A T A T G A T T T A A G A A A G C C C C A
          3850                3860                3870

ASN  GLU  LEU  PHE  ALA  GLY  THR  SER  TYR  ILE  T
A T G A A C T G T T T G C A G G C A C A A G C T A C A T C A
          3880                3890                3900
```

FIG.2V

```
HR  GLU  SER  CYS  LEU  ALA  SER  ASP  ASP  PRO
CCGAAAGCTGTTTGGCAAGTGATGACCCAA
          3910          3920          3930
                              LYS  SER  CYS  VAL  GLN  TYR  PRO  TYR  VAL  TYR  T
                              AAAGCTGCGTACAATACCCTTATGTCTACA
                                        3940          3950          3960

HR  LYS  ALA  ARG  PRO  ASP  GLY  ILE  GLY  ASN
CCAAAGCCCGACCAGATGGCATCGGCAATC
          3970          3980          3990
                              ARG  ASN  PHE  SER  GLU  LEU  SER  ASP  ALA  GLU  L
                              GCAATTTTTCTGAGTTAAGCGATGCTGAAA
                                        4000          4010          4020

YS  ALA  GLN  TYR  LEU  ALA  SER  THR  HIS  PRO
AAGCACAATATTTGGCATCCACGCACCCCC
          4030          4040          4050
                              HIS  GLU  VAL  VAL  SER  ALA  LYS  ASP  TYR  THR  G
                              ATGAGGTTGTCTCTGCCAAAGATTATACAG
                                        4060          4070          4080
```

FIG.2W

```
LY   ILE   TYR   ARG   LEU   LEU   PRO   ASP   PRO   MET
G C A T T T A T C G G T T G T T A C C T G A C C C C A T G G
              4090                4100                4110
                                  ASP   TYR   ARG   SER   ASP   SER   TYR   LEU   ALA   ARG   L
                                  A C T A T C G T T C A G A C T C G T A T T T G G C A C G C C
                                                      4120                4130                4140

EU   ASN   ILE   LYS   ILE   THR   PRO   ASN   LEU   VAL
T T A A C A T C A A A A T C A C C C C A A A T C T G G T C A
              4150                4160                4170
                                  SER   LYS   LEU   LEU   LEU   GLU   ASP   THR   LYS   GLN   T
                                  G T A A A C T G T T A T T A G A A G A C A C C A A G C A A A
                                                      4180                4190                4200

HR   TYR   ASN   ILE   ARG   ASP   MET   ARG   HIS   CYS
C A T A C A A C A T T C G T G A T A T G C G T C A T T G T A
              4210                4220                4230
                                  SER   TYR   HIS   GLY   ALA   ARG   LEU   GLY   ASN   ASP   G
                                  G T T A C C A T G G G G C A A G A T T G G G C A A T G A T G
                                                      4240                4250                4260
```

FIG.2X

```
LY   LYS   PRO   ALA   ASN   GLY   GLY   SER   ILE   VAL
G T A A G C C T G C C A A T G G T G G C T C C A T T G T T C
             4270                4280                4290
                                 LEU   CYS   ASP   ASP   TYR   GLN   GLU   TYR   LEU   ASN   A
                                  T T T G C G A T G A T T A T C A A G A G T A T C T A A A C G
                                                      4300                4310                4320

LA   ASN   ASP   ALA   SER   GLN   ALA   LEU   PHE   ARG
C C A A T G A C G C A T C A C A A G C A T T A T T T A G A C
             4330                4340                4350
                                 PRO   GLY   ALA   ASN   ASP   ALA   PRO   ILE   PRO   LYS   L
                                  C A G G T G C T A A T G A T G C C C C C A T T C C A A A A C
                                                      4360                4370                4380

EU   ALA   TYR   ALA   ARG   SER   SER   VAL   PHE   ASN
T G G C T T A T G C C A G A A G C A G T G T G T T T A A C C
             4390                4400                4410
                                 GLN   GLU   HIS   GLY   LYS   THR   ARG   TYR   GLY   LEU   S
                                  A A G A G C A T G G C A A A A C T C G C T A T G G G T T A A
                                                      4420                4430                4440
```

FIG.2Y

```
ER   PHE  GLU  PHE  LYS  PRO  ASP  THR  PRO  TRP
G T T T T G A G T T T A A G C C T G A C A C G C C A T G G T
            4450              4460              4470
                              PHE  LYS  GLN  ALA  LYS  LEU  ASN  LEU  HIS  GLN  G
                               T T A A G C A A G C A A A A T T A A A C C T A C A C C A A C
                                                 4480              4490              4500

LN   ASN  ILE  GLN  ILE  ILE  ASN  HIS  ASP  ILE
A A A A T A T C C A A A T C A T T A A C C A T G A C A T T A
            4510              4520              4530
                              LYS  LYS  SER  CYS  SER  GLN  TYR  PRO  LYS  VAL  A
                               A A A A A T C G T G C A G C C A A T A T C C T A A G G T G G
                                                 4540              4550              4560

SP   LEU  ASN  CYS  GLY  ILE  SER  GLU  ILE  GLY
A T T T A A A T T G T G G C A T C A G T G A A A T T G G G C
            4570              4580              4590
                              HIS  TYR  GLU  TYR  GLN  ASN  ASN  TYR  ARG  TYR  L
                               A T T A T G A A T A T C A A A A T A A T T A C C G T T A T A
                                                 4600              4610              4620
```

FIG.2Z

YS GLU GLY ARG ALA SER LEU THR GLY LYS
A A G A A G G G C G T G C C A G C T T G A C A G G C A A A C
4630 4640 4650

LEU ASP PHE ASN PHE ASP LEU LEU GLY GLN H
T T G A T T T T A A T T T T G A C C T G C T G G G T C A G C
4660 4670 4680

IS ASP LEU THR VAL LEU ALA GLY ALA ASP
A C G A T T T G A C G G T G T T G G C T G G T G C A G A T A
4690 4700 4710

LYS VAL LYS SER GLN PHE ARG ALA ASN ASN P
A A G T T A A A A G C C A A T T T C G T G C C A A C A A C C
4720 4730 4740

RO ARG ARG THR ILE ILE ASP THR THR GLN
C C A G A C G C A C A A T C A T T G A C A C C A C C C A A G
4750 4760 4770

GLY ASP ALA ILE ILE ASP GLU SER THR LEU T
G C G A T G C C A T C A T T G A T G A A A G C A C G C T G A
4780 4790 4800

FIG.2A'

```
HR   ALA   GLN   GLU   GLN   ALA   LYS   PHE   LYS   GLN
C A G C A C A G G A G C A A G C C A A A T T T A A G C A A T
              4810                4820                4830
                                  SER   GLY   ALA   ALA   TRP   ILE   VAL   LYS   ASN   ARG   L
                                  C G G G G G C G G C A T G G A T T G T C A A A A A T C G C C
                                                4840                4850                4860

EU   GLY   ARG   LEU   GLU   GLU   LYS   ASP   ALA   CYS
T T G G A C G C T T A G A A G A A A A A G A C G C C T G T G
              4870                4880                4890
                                  GLY   ASN   ALA   ASN   GLU   CYS   GLU   ARG   ALA   PRO   I
                                  G C A A T G C C A A T G A A T G T G A A C G C G C C C C C A
                                                4900                4910                4920

LE   HIS   GLY   SER   ASN   GLN   TYR   VAL   GLY   ILE
T T C A T G G C A G T A A C C A A T A T G T G G G C A T T A
              4930                4940                4950
                                  ASN   ASN   LEU   TYR   THR   PRO   ASN   ASP   TYR   VAL   A
                                  A C A A C C T T T A T A C A C C A A A T G A T T A T G T G G
                                                4960                4970                4980
```

FIG.2B'

```
SP   LEU   SER   PHE   GLY   GLY   ARG   LEU   ASP   LYS
A T T T A A G T T T T G G T G G A C G C T T G G A T A A A C
              4990                5000                5010
                                  GLN   ARG   ILE   HIS   SER   THR   ASP   SER   ASN   ILE   I
                                   A A C G C A T T C A C A G C A C C G A T T C A A A C A T C A
                                                      5020                5030                5040

LE   SER   LYS   THR   TYR   THR   ASN   LYS   SER   TYR
T C A G C A A A A C T T A C A C C A A C A A A A G C T A T A
              5050                5060                5070
                                  ASN   PHE   GLY   ALA   ALA   VAL   HIS   LEU   THR   PRO   A
                                   A T T T T G G A G C G G C G G T T C A T C T G A C A C C T G
                                                      5080                5090                5100

SP   PHE   SER   LEU   LEU   TYR   LYS   THR   ALA   LYS
A T T T T A G C C T G T T G T A T A A A A C T G C C A A A G
              5110                5120                5130
                                  GLY   PHE   ARG   THR   PRO   SER   PHE   TYR   GLU   LEU   T
                                   G C T T T C G T A C G C C A A G T T T T T A T G A A C T G T
                                                      5140                5150                5160
```

FIG.2C'

```
YR   ASN   TYR   ASN   SER   THR   ALA   ALA   GLN   HIS
A C A A C T A T A A C A G C A C C G C C G C C C A G C A T A
             5170                5180                5190
                                LYS   ASN   ASP   PRO   ASP   VAL   SER   PHE   PRO   LYS   A
                                 A A A A T G A C C C T G A T G T G T C T T T T C C C A A A C
                                                    5200                5210                5220

RG   ALA   VAL   ASP   VAL   LYS   PRO   GLU   THR   SER
G A G C G G T T G A T G T C A A A C C T G A A A C T T C C A
             5230                5240                5250
                                ASN   THR   ASN   GLU   TYR   GLY   PHE   ARG   TYR   GLN   H
                                 A T A C C A A T G A A T A C G G C T T T C G C T A T C A G C
                                                    5260                5270                5280

IS   PRO   TRP   GLY   ASP   VAL   GLU   MET   SER   MET
A C C C T T G G G G G G A T G T T G A G A T G A G C A T G T
             5290                5300                5310
                                PHE   LYS   SER   ARG   TYR   LYS   ASP   MET   LEU   ASP   L
                                 T C A A A A G C C G T T A C A A G G A C A T G T T A G A T A
                                                    5320                5330                5340
```

FIG.2D'

```
YS   ALA   ILE   PRO   ASN   LEU   THR   LYS   ALA   GLN
A A G C C A T A C C G A A C C T A A C C A A A G C C C A A C
            5350                5360                5370
                                GLN   GLU   TYR   CYS   LYS   ALA   HIS   LEU   ASP   SER   A
                                 A A G A G T A T T G T A A G G C T C A T T T G G A T T C C A
                                              5380                5390                5400

SN   GLU   CYS   VAL   GLY   ASN   PRO   PRO   THR   PRO
A T G A A T G T G T T G G C A A T C C G C C C A C G C C C A
            5410                5420                5430
                                LYS   THR   SER   ASP   GLU   VAL   PHE   ALA   ASN   LEU   T
                                 A A A C C A G T G A T G A G G T A T T T G C C A A C T T A T
                                              5440                5450                5460

YR   ASN   ALA   THR   ILE   LYS   GLY   VAL   SER   VAL
A T A A T G C C A C C A T C A A A G G G G T G A G T G T C A
            5470                5480                5490
                                LYS   GLY   LYS   LEU   ASP   LEU   HIS   ALA   MET   THR   S
                                 A A G G C A A A C T G G A T T T G C A T G C C A T G A C A T
                                              5500                5510                5520
```

FIG.2E'

```
ER  LYS  LEU  PRO  ASP  GLY  LEU  GLU  MET  THR
C A A A A C T G C C A G A T G G T C T T G A A A T G A C C T
          5530                 5540                 5550

LEU  GLY  TYR  GLY  HIS  THR  LYS  LEU  GLY  LYS  P
T G G G T T A T G G T C A T A C C A A A T T G G G G A A A T
          5560                 5570                 5580

HE  ASP  TYR  ILE  ALA  PRO  LYS  ASP  ALA  ASP
T T G A T T A C A T T G C A C C C C A A A G A T G C C C G A T G
          5590                 5600                 5610

GLY  TRP  TYR  GLN  ALA  ARG  PRO  ALA  PHE  TRP  A
G T T G G T A T C A G G C T C G C C C C T T T T T G G G
          5620                 5630                 5640

SP  ALA  ILE  THR  PRO  ALA  ARG  TYR  VAL  VAL
A T G C C C A T C A C C C C C A G C G C C T A T G T G G T C G
          5650                 5660                 5670

GLY  LEU  ASN  TYR  ASP  HIS  PRO  SER  GLN  VAL  T
G T C T T A A A C T A T G A C C C C C C A G T C A A G T A T
          5680                 5690                 5700
```

FIG.2F'

RP GLY ILE GLY THR THR LEU THR HIS SER
G G G G C A T T G G C A C A A C T T T A A C G C A C A G C A
5710 5720 5730
LYS GLN LYS ASP GLU ASN GLU LEU SER ALA L
A A C A A A A A G A T G A A A A T G A G C T A A G T G C C C
5740 5750 5760

EU ARG ILE ARG ASN GLY LYS ARG GLU THR
T T A G A A T C C G A A A T G G C A A A A G A G A A A C A C
5770 5780 5790
GLN THR LEU THR HIS THR ILE PRO LYS ALA T
A A A C C T T A A C G C A C A C A A T A C C C A A A G C C T
5800 5810 5820

YR THR LEU LEU ASP MET THR GLY TYR TYR
A T A C C T T A C T G G A C A T G A C A G G C T A T T A T A
5830 5840 5850
SER PRO THR GLU SER ILE THR ALA ARG LEU G
G C C C A A C T G A G A G C A T C A C C G C T C G T C T T G
5860 5870 5880

FIG.2G'

```
LY   ILE   ASN   ASN   VAL   LEU   ASN   THR   ARG   TYR
G T A T C A A C A A T G T A T T A A A C A C C C G C T A C A
             5890                5900                5910
                                THR   THR   TRP   GLU   ALA   ALA   ARG   GLN   LEU   PRO   S
                                C C A C A T G G G A A G C G G C A C G C C A A C T G C C C A
                                              5920                5930                5940

ER   GLU   ALA   ALA   SER   SER   THR   GLN   SER   THR
G C G A A G C T G C A A G C A G T A C C C A A T C A A C C C
             5950                5960                5970
                                ARG   TYR   ILE   ALA   PRO   GLY   ARG   SER   TYR   PHE   A
                                G T T A C A T T G C A C C A G G T C G C A G T T A C T T T G
                                              5980                5990                6000

                                                  ORF3
LA   SER   LEU   GLU   MET   LYS   PHE   ***     MET   THR
C C A G T C T T G A A A T G A A G T T T T A A T A T G A C C
             6010                6020                6030
                                 CYS   LEU   PRO   LYS   THR   ASN   PRO   ALA   LEU . LYS
                                 T G T T T A C C A A A G A C C A A C C C T G C T T T A A A A
                                              6040                6050                6060
```

FIG.2H'

```
VAL  LYS  HIS  ARG  PHE  LEU  LYS  GLN  VAL  LEU
G T C A A G C A C A G A T T T T A A A G C A G G T G C T G
                6070                 6080                 6090

LEU  LEU  LEU  CYS  VAL  ASP  THR  LEU  THR  ALA
T T A T T G C T T T G T T G A T A C A T T A A C A G C A
                6100                 6110                 6120

GLN  ALA  TYR  ALA  HIS  SER  HIS  HIS  THR  PRO
C A G G C G T A C G C C C A C A G C C A T C A T A C G C C C
                6130                 6140                 6150

ILE  HIS  THR  PRO  THR  HIS  GLU  LEU  PRO  SER
A T T C A T A C A C C C A C G C A T G A G C T G C C A T C T
                6160                 6170                 6180

ALA  ASP  ALA  LEU  SER  ASP  GLU  GLY  LEU  GLY
G C T G A T G C T T T A T C A G A T G A A G G C T T G G G T
                6190                 6200                 6210

LYS  ASP  LEU  GLY  SER  LEU  ASP  SER  LEU  ASP
A A G G A T T T G G G C A G T T T G G A C A G T T T G G A T
                6220                 6230                 6240
```

FIG.2I'

```
SER  PRO  ASP  GLY  LEU  GLY  ASP  GLY  LEU  GLY
AGCCCCAGATGGGTTTGGGTTAGGC
          6250
                    6260
                         ASP  GLY  LEU  GLY  ASP  GLY  LEU  LYS  SER  ASP
                         GATGGGTTTGGGTGATGGCTTAAAAGTGAT
                                   6280
                                             6290
                                                  6300

LYS  ALA  PRO  LEU  PRO  ILE  ASN  ALA  LEU  THR
AAAGCCCCCTTTACCCAATCAACGCCCTTGACC
          6310
                    6320
                         ALA  HIS  GLN  THR  ASN  GLU  SER  GLN  PRO  ALA
                         GCCCATCAGACCAATGAGAGCCCAGCCTGCC
                                   6340
                                             6350
                                                  6360

PRO  PRO  SER  VAL  ASP  VAL  ASN  PHE  LEU  LEU
CCACCCGAGCCGTAGATGTCAATTTTTTACTT
          6370
                    6380
                         ALA  GLN  PRO  GLU  ALA  PHE  TYR  HIS  VAL  PHE
                         GCCCAGCCAGAGGCCATTTTATCATGTCTTT
                                   6400
                                             6410
                                                  6420
```

FIG.2J'

```
 HIS  GLN  ALA  ILE  VAL  GLN  ASP  ASP  VAL  ALA
C A T C A A G C G A T T G T G C A A G A T G A T G T G G C A
          6430           6440           6450
                          THR  LEU  ARG  LEU  LEU  LEU  PRO  PHE  TYR  ASP
                         A C A T T A C G C T T G T T A T T G C C A T T T T A T G A C
                                    6460           6470           6480

 ARG  LEU  PRO  ASP  ASP  TYR  GLN  ASP  ASP  VAL
C G C C T G C C T G A T G A T T A T C A A G A T G A T G T T
          6490           6500           6510
                          LEU  LEU  LEU  PHE  ALA  GLN  SER  LYS  LEU  ALA
                         T T G T T G T T A T T T G C C C A A A G T A A A C T T G C C
                                    6520           6530           6540

 LEU  SER  ASP  GLY  ASN  THR  LYS  LEU  ALA  LEU
C T A A G T G A T G G C A A T A C C A A A T T G G C A T T G
          6550           6560           6570
                          ASN  LEU  LEU  THR  ASP  LEU  SER  ASN  LYS  GLU
                         A A T C T G C T G A C C G A T T T G A G T A A C A A A G A G
                                    6580           6590           6600
```

FIG.2K'

```
 PRO  THR  LEU  THR  ALA  VAL  LYS  LEU  GLN  LEU
C C A A C A C T T A C G G C G G T A A A A T T A C A A C T T
              6610                 6620                 6630
                                    ALA  SER  LEU  LEU  LEU  THR  ASN  LYS  HIS  ASP
                                   G C T T C C T T G T T G C T G A C C A A C A A G C A C G A T
                                                        6640                 6650                 6660

 LYS  HIS  ALA  GLN  MET  VAL  LEU  ASP  GLU  LEU
A A A C A C G C C C A A A T G G T G C T A G A T G A A C T C
              6670                 6680                 6690
                                    LYS  ASP  ASP  ALA  HIS  PHE  LEU  LYS  LEU  SER
                                   A A A G A T G A T G C C C A C T T T T T A A A A T T A A G C
                                                        6700                 6710                 6720

 LYS  LYS  GLU  GLN  ARG  TRP  VAL  LEU  SER  GLN
A A A A A A G A G C A A A G A T G G G T G C T A T C G C A A
              6730                 6740                 6750
                                    SER  ARG  TYR  LEU  HIS  LYS  LYS  TYR  LYS  MET
                                   A G T C G C T A T T T A C A T A A A A A A T A T A A A A T G
                                                        6760                 6770                 6780
```

FIG.2L'

```
 GLY   LEU   ASP   LEU   GLY   ILE   ASN   TYR   LEU   HIS
G G C T T G G A T T T G G G C A T C A A C T A T C T G C A T
            6790                6800                6810
                                 LEU   ASP   ASN   ILE   ASN   ALA   ALA   SER   THR   ILE
                                T T G G A T A A T A T C A A C G C C G C C T C C A C C A T C
                                            6820                6830                6840

 THR   GLN   PRO   ASN   ILE   LYS   LYS   ASP   ALA   PRO
A C C C A G C C C A A T A T T A A A A A A G A T G C C C C A
            6850                6860                6870
                                 LYS   PRO   ALA   HIS   GLY   LEU   ALA   LEU   SER   LEU
                                A A A C C T G C T C A T G G G C T T G C C T T A T C G C T T
                                            6880                6890                6900

 GLY   VAL   ASN   LYS   TYR   THR   PRO   LEU   SER   HIS
G G T G T G A A T A A A T A C A C G C C G C T T A G T C A T
            6910                6920                6930
                                 GLY   MET   SER   ILE   TYR   THR   ALA   LEU   ASP   VAL
                                G G C A T G A G T A T T T A T A C A G C C C T A G A T G T T
                                            6940                6950                6960
```

FIG.2M'

ASP GLY LYS PHE TYR ASP ASP LYS SER HIS
G A T G G T A A A T T T T A T G A T G A C A A A A G C C A C
6970 6980 6990

ASN GLU LEU ALA VAL PHE ALA HIS ALA GLY
A A T G A A C T G G C G G T T T T T G C T C A T G C T G G A
7000 7010 7020

LEU ARG LYS ASP HIS GLN LYS GLY TYR VAL
C T A A G A A A A G A T C A C C A A A A A G G T T A T G T T
7030 7040 7050

ASP VAL VAL PRO PHE VAL GLY ARG ILE PHE
G A T G T C G T A C C T T T T G T T G G G C G T A T T T T T
7060 7070 7080

ALA THR ASN GLN GLN HIS GLY ARG LEU SER
G C C A C C A A T C A G C A G C A T G G C A G A T T A T C C
7090 7100 7110

PRO ARG LYS ASP SER GLN GLY VAL ALA PHE
C C C A G A A A A G A C A G T C A G G G C G T G G C G T T T
7120 7130 7140

FIG.2N'

```
 GLY  SER  HIS  HIS  ARG  ILE  ASN  ASP  LYS  TRP
G G C A G C C A T C A T C G G A T C A A T G A T A A A T G G
            7150              7160              7170
                                GLN  ASN  ALA  PHE  PHE  ALA  ARG  MET  GLU  LYS
                               C A A A A T G C G T T T T T G C A C G C A T G G A A A A A
                                                 7180              7190              7200

 GLY  ASN  TYR  THR  GLU  ARG  TYR  GLN  GLY  TYR
G G C A A T T A T A C C G A G C G T T A T C A A G G T T A T
            7210              7220              7230
                                ASP  GLY  LYS  ARG  TYR  HIS  VAL  ASN  ASP  THR
                               G A T G G C A A G C G T T A T C A T G T G A A T G A C A C C
                                                 7240              7250              7260

 ILE  LEU  LEU  GLN  ASP  GLY  PRO  ASN  ARG  ARG
A T T T T G T T G C A A G A T G G C C C A A A T C G T C G T
            7270              7280              7290
                                TYR  SER  LEU  GLY  VAL  GLY  TYR  GLN  LEU  SER
                               T A C T C T T T G G G C G T G G G G T A T C A G C T T A G C
                                                 7300              7310              7320
```

FIG.20'

```
 HIS  LEU  GLN  ASP  ALA  THR  LYS  SER  SER  HIS
C A T C T G C A A G A T G C A A C A A A A A G C A G T C A T
            7330              7340              7350
                               ALA  THR  LYS  ILE  HIS  PHE  GLY  VAL  LEU  GLN
                              G C C A C A A A G A T A C A T T T G G G G T G T T G C A A
                                                7360              7370              7380

 ARG  LEU  PRO  ASN  GLY  LEU  THR  VAL  GLN  GLY
A G A T T G C C A A A T G G T C T G A C C G T G C A A G G T
            7390              7400              7410
                               ARG  VAL  SER  ALA  GLU  ARG  GLU  ARG  TYR  HIS
                              A G A G T G A G T G C T G A G C G T G A G C G T T A T C A T
                                                7420              7430              7440

 GLY  LYS  LEU  LEU  ARG  LEU  VAL  ASN  PRO  ASP
G G T A A A T T A T T G C G T C T G G T T A A T C C T G A T
            7450              7460              7470
                               ASP  VAL  TYR  ARG  THR  ASP  LYS  THR  LEU  THR
                              G A T G T G T A T C G C A C A G A T A A A A C C C T A A C C
                                                7480              7490              7500
```

FIG.2P'

```
 LEU  GLN  THR  SER  ILE  TRP  HIS  LYS  ASP  ILE
C T A C A A A C C T C C A T T G G C A C A A A G A C A T T
                7510
                          7520
                                    7530
          HIS  TRP  LEU  GLY  LEU  THR  PRO  LYS  LEU  THR
C A C T G G C T T G G A T T A A C G C C A A A G C T G A C T
                    7540
                              7550
                                        7560

 TYR  ARG  TYR  SER  LYS  ASN  ASN  SER  ASN  LEU
T A T C G T T A C A G T A A A A A T A A C A G T A A C T T A
                7570
                          7580
                                    7590
          PRO  ALA  LEU  TYR  SER  HIS  ASN  LYS  GLN  ASN
C C A G C A C T T T A T A G C C C A T A A C A A A A C A A T
                    7600
                              7610
                                        7620

 PHE  TYR  LEU  GLU  LEU  GLY  ARG  SER  PHE  ***
T T T A T T T G G A G C T T G G T C G G T T T T A A
                7630
                          7640
                                    7650
```

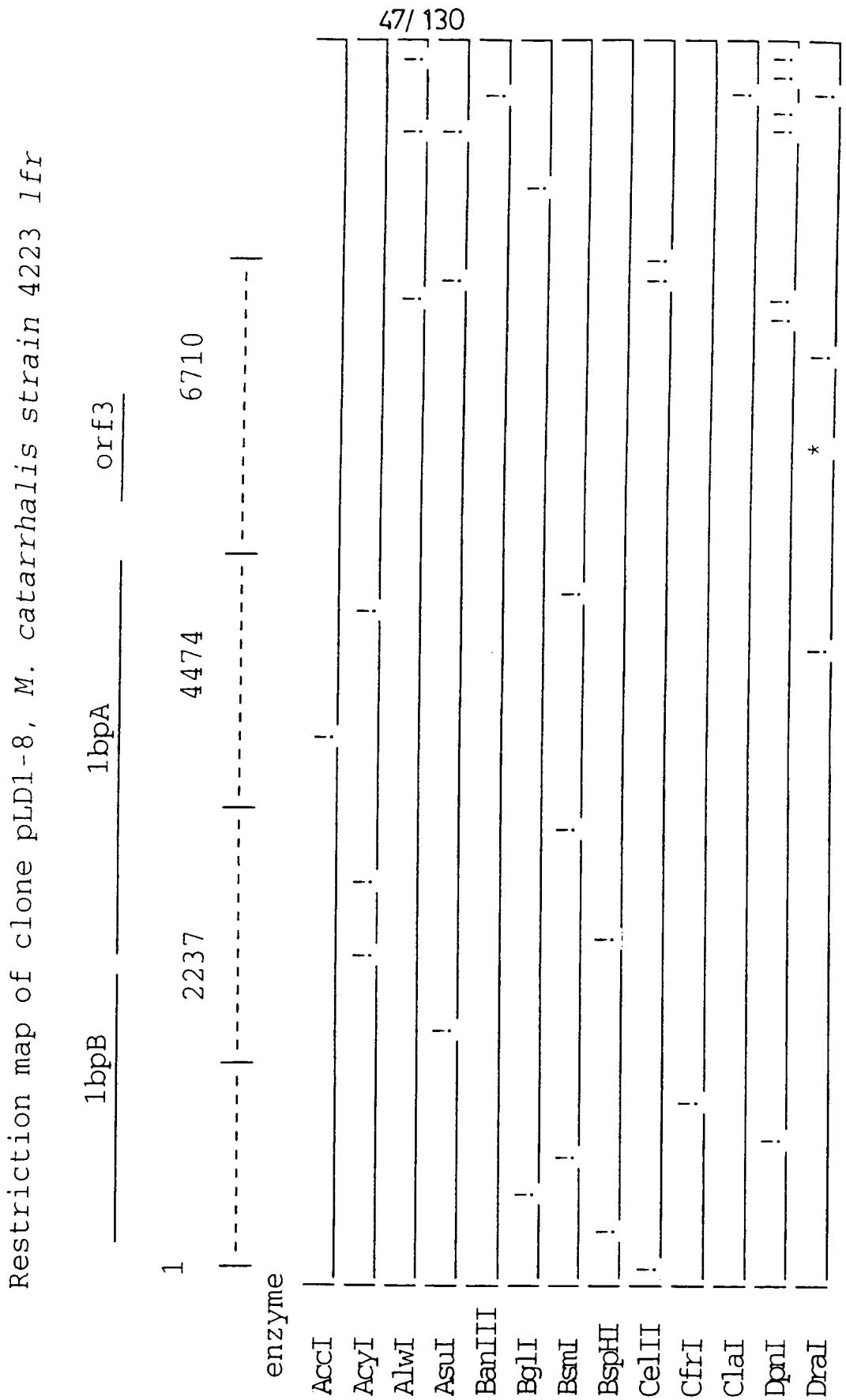


FIG.3A

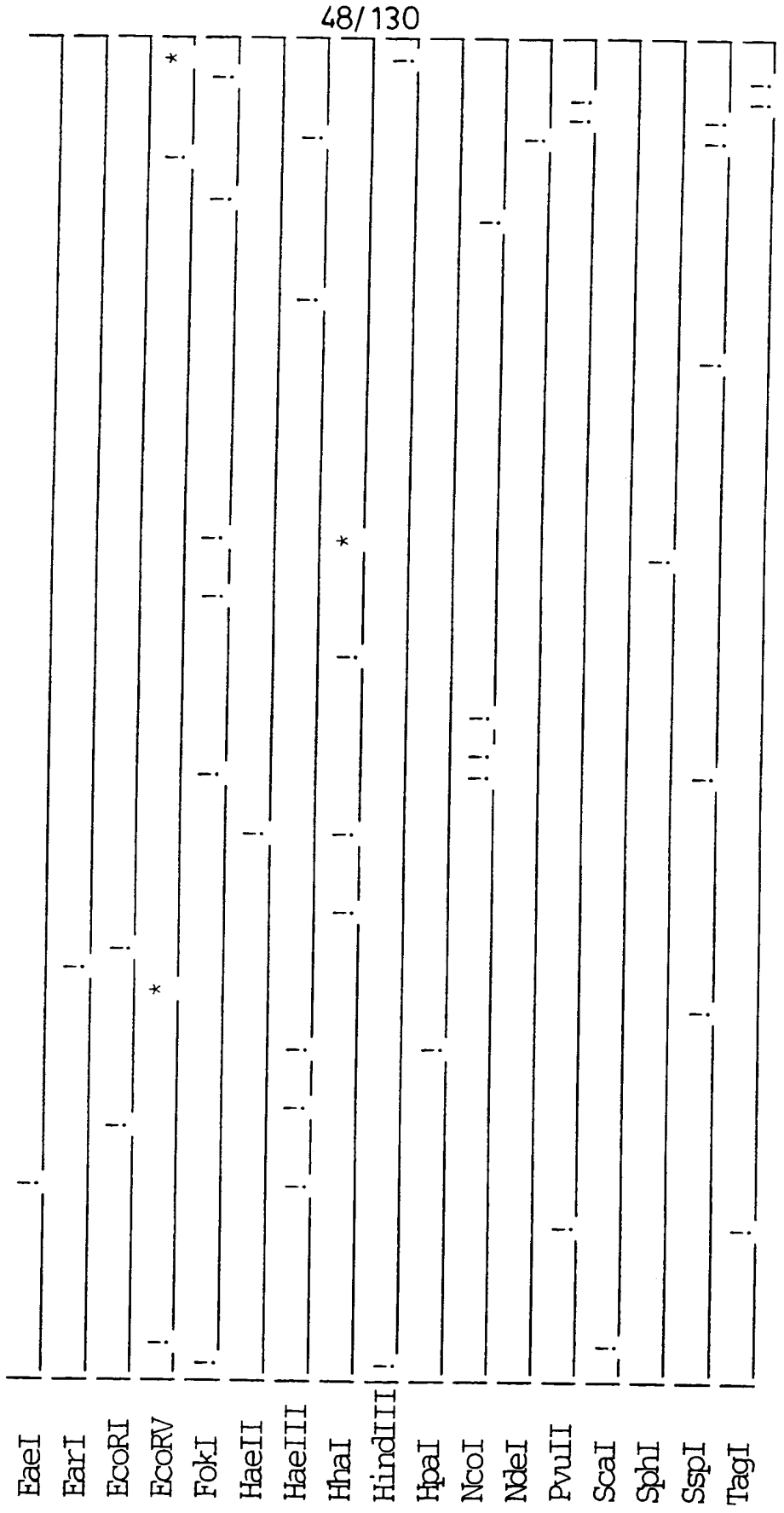


FIG.3B

FIG.4A *M. catarrhalis* Q8 *lfr* sequence

```
A A G C T T A G C A T G A T G G C A T C G G C T G A T T G T
              10                  20                  30
                  C T T T T T G C C T T G T T G T G T G T T T G T G G G A G T
                                      40                  50                  60

T G A T T G T A C T T A C C T T A G T G G T G G A T G C T T
              70                  80                  90
                                      -35                             -10
                  G G G C T G A T T T A A T T A A A T T T A A T C A A A G C G
                                      100                 110                 120

                                  RBS
G T C T T C A C A A C A C A C C A A A C G A G A T A T C A C
              130                 140                 150
                                  Lbp2
                                  MET  SER  THR  VAL  LYS  THR  PRO  HIS  ILE  PHE
                  C A T G A G T A C T G T C A A A A C C C C C C A T A T T T T
                                      160                 170                 180

   TYR  GLN  LYS  ARG  THR  LEU  SER  LEU  ALA  ILE
C T A C C A A A A A C G C A C C C T T A G C C T T G C C A T
              190                 200                 210
                                  ALA  SER  ILE  PHE  ALA  ALA  LEU  VAL  MET  THR
                  C G C C A G T A T T T T T G C T G C C T T G G T G A T G A C
                                      220                 230                 240
```

FIG.4B

```
   GLY  CYS  ARG  SER  ASP  ASP  ILE  SER  VAL  ASN
A G G C T G C C G C T C T G A T G A C A T C A G C G T C A A
              250                  260                  270
                    ALA  PRO  ASN  VAL  THR  GLN  LEU  PRO  GLN  GLY
                  T G C A C C C A A T G T T A C C C A G C T G C C C C A A G G
                                    280                  290                  300

   THR  VAL  SER  PRO  THR  PRO  ASN  THR  GLY  HIS
C A C G G T T T C A C C A A C G C C G A A C A C A G G T C A
              310                  320                  330
                    ASP  ASN  ALA  ASN  ASN  THR  ASN  ASN  GLN  GLY
                  T G A C A A C G C C A A T A A C A C C A A C A A T C A G G G
                                    340                  350                  360

   ASN  ASN  THR  ASP  ASN  SER  THR  SER  THR  THR
C A A C A A C A C G G A T A A C A G C A C C A G C A C A A C
              370                  380                  390
                    ASP  PRO  ASN  GLY  ASP  ASN  ASN  GLN  LEU  THR
                  T G A C C C A A A T G G C G A T A A C A A C C A A C T G A C
                                    400                  410                  420
```

FIG.4C

```
   GLN  ALA  GLN  LYS  THR  ALA  ALA  ALA  ALA  GLY
A C A A G C G C A A A A A A C T G C C G C C G C C G C A G G
              430                  440                  450
                    PHE  PHE  VAL  MET  GLY  LYS  ILE  ARG  ASP  THR
                  G T T T T T T G T G A T G G G T A A A A T T C G T G A T A C
                                   460                  470                  480

   SER  GLU  LYS  ASN  ASP  PRO  ASP  TYR  SER  ASP
C A G C G A A A A A A A T G A C C C A G A T T A T A G T G A T
              490                  500                  510
                    ASP  LEU  LYS  GLN  GLN  TRP  LEU  GLY  LYS  LEU
                  T G A T T T A A A A C A G C A G T G G C T G G G C A A A T T
                                   520                  530                  540

   TYR  VAL  GLY  ILE  ASP  ALA  HIS  ARG  PRO  ASP
A T A T G T T G G T A T T G A T G C C C A T C G C C C A G A
              550                  560                  570
                    GLY  ILE  GLY  LYS  GLY  LYS  ASN  LEU  ARG  GLN
                  T G G C A T C G G A A A A G G T A A A A A C T T G C G T C A
                                   580                  590                  600
```

FIG.4D

```
  PRO  ILE  THR  ALA  ASN  ASP  ILE  LYS  PRO  LEU
G C C C A T C A C C G C C A A T G A C A T C A A A C C C T T
      610       620       630       640       650

  TYR  PHE  ASN  LYS  PHE  PRO  ALA  LEU  SER  ASP
G T A T T T T A A C A A A T T C C C C T G C A T T G T C T G A
      660       670       680       690       700

  LEU  HIS  LEU  ASP  SER  GLU  ARG  HIS  ARG  PHE
T T T G C A C T T A G A C A G T G A A C G C C A T C G T T T T
      710       720       730       740       750

  ASP  PRO  GLN  LYS  ILE  ASN  THR  ILE  LYS  VAL
T G A C C C C C C A A A A G A T A A A C A C C A T T A A A G T
      760       770       780       790       800

  TYR  GLY  TYR  GLY  ASN  LEU  THR  THR  PRO  SER
G T A T G G G T T A T G G T A A C T T A A C A A C C A T C
      810       820       830       840       850

  ASN  ASN  ASN  THR  HIS  ILE  ASN  HIS  GLN  GLN
C A A C A A C A A C A C T C A C A T C A A T C A T C A G C A
      860       870       880       890       900
```

FIG.4E

```
  ALA  ASP  ASN  LYS  LYS  ASN  ASN  LYS  PRO  VAL
A G C T G A T A A T A A G A A A A A T A A C A A G C C C T G T
      790       800       810       820       830       840

  ASP  PRO  TYR  GLU  ASN  ILE  ARG  PHE  GLY  TYR
T G A C C C C T T A T G A A A A A T A T C C G T T T T G G G T A
      790       800       810       820       830       840

  LEU  GLU  LEU  GLN  GLY  SER  SER  LEU  THR  GLN
T C T T G A A C T A C A A G G A A G C A G C C T G A C C C C A
      850       860       870       880       890       900

  LYS  ASN  ALA  ASP  ASN  GLN  ASN  GLU  GLN  ASP
A A A A A A A T G C C C G A T A A T C A A A A A T G A G C A A G A
      850       860       870       880       890       900

  ARG  ILE  PRO  LYS  PRO  MET  PRO  ILE  LEU  PHE
C C G C A T T C C C C A A A C C C C A T G C C C C A T T T T G T T
      910       920       930       940       950       960

  TYR  HIS  GLY  GLU  ASN  ALA  SER  SER  GLN  LEU
T T A T C A T G G A G A A A A A C G C C C A G C C A G C C T
      910       920       930       940       950       960
```

FIG.4F

```
   PRO  SER  ALA  GLY  LYS  PHE  ASN  TYR  THR  GLY
G C C C A G C G C T G G T A A A T T T A A C T A C A C A G G
              970                  980                  990
                    ASN  TRP  LEU  TYR  LEU  SER  ASP  VAL  LYS  LYS
                  C A A C T G G C T G T A C C T A A G T G A T G T C A A A A A
                                      1000                 1010                 1020

   ARG  PRO  ALA  LEU  SER  ALA  SER  ASP  GLU  ARG
A C G C C C T G C C C T T T C A G C A T C A G A T G A G C G
              1030                 1040                 1050
                    VAL  GLY  VAL  TYR  LEU  ASN  ALA  SER  GLY  LYS
                  A G T G G G G G T C T A T C T C A A T G C C A G T G G C A A
                                      1060                 1070                 1080

   ALA  ASN  GLU  GLY  ASP  VAL  VAL  SER  ALA  ALA
A G C C A A C G A G G G C G A T G T C G T C A G T G C C G C
              1090                 1100                 1110
                    HIS  ILE  TYR  LEU  ASN  GLY  PHE  GLN  TYR  LYS
                  C C A C A T T T A T C T A A A C G G C T T T C A A T A T A A
                                      1120                 1130                 1140
```

FIG.4G

```
   HIS  THR  PRO  ALA  THR  TYR  GLN  VAL  ASP  PHE
G C A C A C G C C T G C C A C T T A T C A G G T G G A T T T
          1150           1160           1170
                ASP  THR  ASN  SER  LEU  THR  GLY  LYS  LEU  SER
              T G A C A C A A A C T C A T T A A C A G G C A A G C T G T C
                          1180           1190           1200

   TYR  TYR  ASP  ASN  PRO  ASN  GLN  GLN  ASN  ASN
C T A T T A T G A C A A T C C C A A T C A G C A A A A T A A
          1210           1220           1230
                LYS  GLY  GLU  TYR  LEU  LYS  SER  GLN  PHE  ASP
              T A A A G G C G A A T A T C T C A A A A G C C A A T T T G A
                          1240           1250           1260

   THR  THR  LYS  LYS  VAL  ASN  GLU  THR  ASP  VAL
C A C T A C C A A A A A A G T C A A T G A A A C C G A T G T
          1270           1280           1290
                TYR  GLN  ILE  ASP  ALA  LYS  ILE  ASN  GLY  ASN
              G T A T C A A A T T G A T G C C A A A A T C A A C G G T A A
                          1300           1310           1320
```

FIG.4H

```
   ARG   PHE   VAL   GLY   THR   ALA   LYS   SER   LEU   VAL
C C G C T T T G T C G G T A C G G C C A A A T C T T T G G T
             1330                1340                1350
                                    ASN   GLU   LYS   THR   GLN   THR   ALA   PRO   PHE   ILE
                                  T A A T G A G A A A A C A C A A A C C G C A C C T T T T A T
                                                1360                1370                1380

   LYS   GLU   LEU   PHE   SER   LYS   LYS   ALA   ASN   PRO
C A A A G A G C T G T T C T C C A A A A A A G C C A A C C C
             1390                1400                1410
                                    ASN   ASN   PRO   ASN   PRO   ASN   SER   ASP   THR   LEU
                                  C A A T A A C C C A A A C C C T A A T T C A G A C A C G C T
                                                1420                1430                1440

   GLU   GLY   GLY   PHE   TYR   GLY   GLU   SER   GLY   ASP
A G A A G G C G G A T T T T A T G G T G A G T C G G G C G A
             1450                1460                1470
                                    GLU   LEU   ALA   GLY   LYS   PHE   LEU   SER   ASN   ASP
                                  T G A G C T G G C G G G T A A A T T T T T A T C C A A T G A
                                                1480                1490                1500
```

FIG.4I

```
   ASN  ALA  SER  TYR  VAL  VAL  PHE  GLY  GLY  LYS
C A A C G C A T C T T A T G T G G T C T T T G G T G G C A A
            1510              1520              1530
                                 ARG  ASP  LYS  THR  THR  LYS  PRO  VAL  ALA  THR
                               A C G A G A C A A A A C G A C T A A A C C T G T C G C C A C
                                                 1540              1550              1560

   LYS  THR  VAL  TYR  PHE  SER  ALA  GLY  PHE  GLU
A A A A A C G G T G T A T T T T A G T G C A G G C T T T G A
            1570              1580              1590
                                 LYS  PRO  SER  THR  SER  PHE  VAL  ASP  ASN  GLU
                               A A A A C C C A G C A C C A G T T T T G T G G A T A A T G A
                                                 1600              1610              1620

   THR  ILE  GLY  GLY  ILE  ILE  ASP  ARG  LYS  GLY
A A C G A T T G G T G G A A T T A T T G A C C G T A A A G G
            1630              1640              1650
                                 LEU  ASN  ASN  HIS  ILE  ASN  GLU  ASP  GLU  ILE
                               G T T A A A T A A T C A C A T T A A T G A A G A T G A A A T
                                                 1660              1670              1680
```

FIG.4J

```
   ILE   PRO   SER   ASP   ASP   SER   TYR   TYR   GLY   TYR
T A T T C C C A G T G A T G A T A G T T A T T A T G G A T A
            1690                1700                1710
                  THR   TRP   GLY   LYS   PRO   GLU   LYS   GLN   PHE   THR
               T A C T T G G G G C A A G C C A G A G A A G C A G T T C A C
                                  1720                1730                1740

   LYS   LYS   VAL   SER   SER   SER   THR   GLN   VAL   VAL
C A A A A A A G T C A G C A G C A G C A C C C A A G T C G T
            1750                1760                1770
                  PRO   ALA   TYR   PHE   GLY   GLN   HIS   ASP   LYS   PHE
               G C C A G C T T A T T T T G G G C A A C A T G A T A A A T T
                                  1780                1790                1800

   TYR   PHE   ASN   GLY   ASN   TYR   TYR   ASP   LEU   SER
T T A T T T T A A T G G C A A C T A T T A T G A C C T A T C
            1810                1820                1830
                  ALA   SER   ARG   VAL   ASP   LYS   LEU   ALA   PRO   ALA
               A G C C A G T C G T G T T G A T A A A T T A G C C C C T G C
                                  1840                1850                1860
```

FIG.4K

```
ASP ALA VAL LYS ALA ASN GLN SER ILE LYS
C G A T G C T G T C A A A G C C A A C C A A T C C A T T A A
1870 1880 1890

GLU LYS TYR PRO ASN ALA THR LEU ASN LYS
A G A A A A A T A C C C T A A T G C C A C A C T A A A T A A
1900 1910 1920

ASP ASN GLN VAL THR ALA ILE VAL LEU GLN
G G A C A A C C A A G T T A C C G C C A T C G T G C T A C A
1930 1940 1950

GLU ALA LYS ASP ASN LYS PRO TYR THR ALA
A G A A G C C A A A G A T A A T A A G C C T T A T A C C G C
1960 1970 1980

ILE ARG ALA LYS SER TYR GLN HIS ILE SER
C A T T C G T G C C A A A A G C T A T C A G C A C A T C A G
1990 2000 2010

PHE GLY GLU THR LEU TYR ASN ASP ALA ASN
T T T T G G C G A G A C G C T G T A T A A C G A T G C C A A
2020 2030 2040
```

FIG.4L

GLN THR PRO THR ARG SER TYR PHE VAL GLN
C C A A A C C C C A A C A C G C A G T T A T T T T G T G C A
2050 2060 2070

GLY GLY ARG ALA ASP THR SER THR THR LEU
A G G C G G T A G G G C A G A T A C C A G C A C A A C T T T
2080 2090 2100

PRO GLN ALA GLY LYS PHE THR TYR ASN GLY
G C C C C A G G C A G G T A A A T T C A C T T A C A A C G G
2110 2120 2130

LEU TRP ALA GLY TYR LEU THR GLN LYS LYS
T C T T T G G G C A G G C T A C C T G A C C C A A A A A A A
2140 2150 2160

ASP LYS GLY TYR SER ASP ASN ALA GLU THR
G G A C A A A G G T T A T A G C G A T A A T G C A G A A A C
2170 2180 2190

ILE LYS GLU LYS GLY HIS PRO GLY TYR LEU
C A T C A A G G A A A A A G G T C A T C C A G G T T A T C T
2200 2210 2220

FIG.4M

```
   LEU  THR  GLU  ASN  PHE  THR  PRO  GLU  ASP  ASP
G T T A A C C G A A A A C T T C A C C C C A G A A G A T G A
            2230              2240              2250
                                 ASP  ASP  ASP  LEU  THR  ALA  SER  ASP  ASP  SER
                               T G A C G A T G A T T T G A C C G C A T C T G A T G A T T C
                                              2260              2270              2280

   GLN  ASP  ASP  ASN  THR  HIS  GLY  ASP  ASP  ASP
A C A A G A T G A T A A T A C A C A T G G C G A T G A T G A
            2290              2300              2310
                                 LEU  ILE  ALA  SER  ASP  ASP  SER  GLN  ASP  ASP
                               T T T G A T T G C A T C T G A T G A T T C A C A A G A T G A
                                              2320              2330              2340

   ASP  ALA  ASP  GLY  ASP  ASP  ASP  SER  ASP  ASP
T G A C G C A G A T G G A G A T G A C G A T T C A G A T G A
            2350              2360              2370
                                 LEU  GLY  ASP  GLY  ALA  ASP  ASP  ASP  ALA  ALA
                               T T T G G G T G A T G G T G C A G A T G A T G A C G C C G C
                                              2380              2390              2400
```

FIG.4N

```
   GLY  LYS  VAL  TYR  HIS  ALA  GLY  ASN  ILE  ARG
A G G C A A A G T G T A T C A T G C A G G T A A T A T T C G
          2410              2420              2430
                                PRO  GLU  PHE  GLU  ASN  LYS  TYR  LEU  PRO  ILE
                    C C C T G A A T T T G A A A A C A A A T A C T T G C C C A T
                                        2440              2450              2460

   ASN  GLU  PRO  THR  HIS  GLU  LYS  THR  PHE  ALA
T A A T G A G C C T A C T C A T G A A A A A A C C T T T G C
          2470              2480              2490
                                LEU  ASP  GLY  LYS  ASN  LYS  ALA  LYS  PHE  GLU
                    C C T A G A T G G T A A A A A T A A A G C T A A G T T T G A
                                        2500              2510              2520

  VAL  ASP  PHE  ASN  THR  ASN  SER  LEU  THR  GLY
A G T G G A T T T T A A C A C C A A C A G C C T A A C T G G
          2530              2540              2550
                                LYS  LEU  ASN  ASP  GLU  ARG  GLY  ASP  ILE  VAL
                    T A A A T T A A A C G A T G A G A G A G G T G A T A T C G T
                                        2560              2570              2580
```

FIG.4O

```
   PHE  ASP  ILE  LYS  ASN  GLY  LYS  ILE  ASP  GLY
C T T T G A T A T C A A A A A T G G C A A A A T T G A T G G
            2590              2600              2610
                  THR  GLY  PHE  THR  ALA  LYS  ALA  ASP  VAL  PRO
                C A C A G G A T T T A C C G C C A A A G C C G A T G T G C C
                                  2620              2630              2640

   ASN  TYR  ARG  GLU  GLU  VAL  GLY  ASN  ASN  GLN
A A A C T A T C G T G A A G A A G T G G G T A A C A A C C A
            2650              2660              2670
                  GLY  GLY  GLY  PHE  LEU  TYR  ASN  ILE  LYS  ASP
                A G G T G G C G G T T T C T T A T A C A A C A T C A A A G A
                                  2680              2690              2700

   ILE  ASP  VAL  LYS  GLY  GLN  PHE  PHE  GLY  THR
T A T T G A T G T T A A G G G G C A A T T T T T T G G C A C
            2710              2720              2730
                  ASN  GLY  GLU  GLU  LEU  ALA  GLY  GLN  LEU  HIS
                A A A T G G C G A A G A G T T G G C A G G A C A G T T A C A
                                  2740              2750              2760
```

FIG.4P

```
   HIS  ASP  LYS  GLY  ASP  GLY  ILE  ASN  ASP  THR
T C A T G A C A A A G G C G A T G G C A T C A A T G A C A C
           2770           2780           2790
                                ALA  GLU  LYS  ALA  GLY  ALA  VAL  PHE  GLY  ALA
                              C G C C G A A A A A G C A G G G G C T G T C T T T G G G G C
                                                2800           2810           2820

   VAL  LYS  ASP  LYS  ***
T G T T A A A G A T A A A T A A A G C C C C C C T T C A T C
           2830           2840           2850
                              A T C G T T T A G T C G C T T G A C C G A C A G T T G A T G
                                                2860           2870           2880

A C G C C C T T G G C A A T G T C T T A A A A C A G C A C T
           2890           2900           2910
                              T T G A A A C A G T G C C T T G G G C G A A T T C T T G G A
                                                2920           2930           2940

T A A A T G C A C C A G A T T T G C C T T G G G C T A A T A
           2950           2960           2970
                                     -10                          -35
                              T C T T G A T A A A A C A T C G C C A T A A A A T A G A A A
                                                2980           2990           3000
```

FIG.4Q

```
                                        Lbp1
                    RBS                 MET  SER  LYS
A T A A A G T T T A G G A T T T T T T A T G T C A A A A
                3010                3020                3030
                                     SER  ILE  THR  LYS  THR  GLN  THR  PRO  SER  VAL  H
                                    T C T A T C A C A A A A A C A C A A A C A C C A T C A G T C C
                                                        3040                3050                3060

IS  THR  MET  THR  THR  HIS  ARG  LEU  ASN  LEU
A T A C C A T G A C C A C G C A C C G C T T A A A C C T T G
                3070                3080                3090
                                     ALA  ILE  LYS  ALA  ALA  LEU  PHE  GLY  VAL  ALA  V
                                    C C A T C A A A G C G G C G T T A T T T G G T G T G G C A G
                                                        3100                3110                3120

AL  LEU  PRO  LEU  SER  VAL  TRP  ALA  GLN  GLU
T T T T A C C C C T A T C C G T C T G G G C G C A A G A G A
                3130                3140                3150
                                     ASN  THR  GLN  THR  ASP  ALA  ASN  SER  ASP  ALA  L
                                    A C A C T C A G A C A G A T G C C A A C T C T G A T G C C A
                                                        3160                3170                3180
```

FIG.4R

```
YS   ASP   THR   LYS   THR   PRO   VAL   VAL   TYR   LEU
A A G A C A C A A A A A C C C C T G T C G T C T A T T T A G
              3190                3200                3210
                                  ASP   ALA   ILE   THR   VAL   THR   ALA   ALA   PRO   SER   A
                                   A T G C C A T C A C G G T A A C C G C C G C C C C A T C T G
                                                       3220                3230                3240

LA   PRO   VAL   SER   ARG   PHE   ASP   THR   ASP   VAL
C C C C T G T T T C T C G G T T T G A C A C C G A T G T A A
              3250                3260                3270
                                  THR   GLY   LEU   GLY   LYS   THR   VAL   LYS   THR   ALA   A
                                   C A G G G C T T G G C A A A A C C G T C A A A A C C G C T G
                                                       3280                3290                3300

SP   THR   LEU   ALA   LYS   GLU   GLN   VAL   GLN   GLY
A C A C G C T G G C A A A A G A A C A A G T A C A G G G C A
              3310                3320                3330
                                  ILE   ARG   ASP   LEU   VAL   ARG   TYR   GLU   THR   GLY   V
                                   T T C G T G A T T T G G T G C G T T A T G A A A C T G G G G
                                                       3340                3350                3360
```

FIG.4S

```
AL  SER  VAL  VAL  GLU  GLN  GLY  ARG  GLY  GLY
T G A G T G T G G T T G A G C A G G G G C G T G G T G G C A
             3370              3380              3390
                               SER  SER  GLY  PHE  ALA  ILE  HIS  GLY  VAL  ASP  L
                                G C A G C G G A T T T G C C A T T C A T G G C G T G G A T A
                                                  3400              3410              3420

YS  ASN  ARG  VAL  GLY  ILE  THR  VAL  ASP  GLY
A A A A C C G A G T G G G C A T T A C C G T A G A T G G C A
             3430              3440              3450
                               ILE  ALA  GLN  ILE  GLN  SER  TYR  LYS  ASP  GLU  S
                                T T G C C C A A A T T C A A T C C T A C A A A G A C G A A T
                                                  3460              3470              3480

ER  THR  LYS  ARG  ALA  GLY  ALA  GLY  SER  GLY
C C A C T A A G C G A G C T G G G G C A G G C T C T G G G G
             3490              3500              3510
                               ALA  MET  ASN  GLU  ILE  GLU  ILE  GLU  ASN  ILE  A
                                C G A T G A A C G A G A T A G A G A T T G A A A A C A T T G
                                                  3520              3530              3540
```

FIG.4T

```
LA  ALA  VAL  ALA  ILE  ASN  LYS  GLY  GLY  ASN
CCGCCGTTGCCATCAATAAAGGCGGTAATG
          3550           3560           3570
                         ALA  LEU  GLU  ALA  GLY  SER  GLY  ALA  LEU  GLY  G
                          CCTTAGAAGCAGGCTCTGGTGCGTTGGGTG
                                    3580           3590           3600

LY  SER  VAL  ALA  PHE  HIS  THR  LYS  ASP  VAL
GTTCGGTGGCGTTTCATACCAAAGATGTGA
          3610           3620           3630
                         SER  ASP  VAL  LEU  LYS  SER  GLY  ASN  ASN  LEU  G
                          GCGATGTCTTAAAATCTGGTAACAATCTTG
                                    3640           3650           3660

LY  ALA  GLN  SER  LYS  THR  THR  TYR  ASN  SER
GTGCTCAAAGCAAAACCACTTATAACAGCA
          3670           3680           3690
                         LYS  ASN  ASP  HIS  PHE  SER  GLN  THR  LEU  ALA  A
                          AAAATGACCATTTTAGTCAGACGCTGGCAG
                                    3700           3710           3720
```

FIG.4U

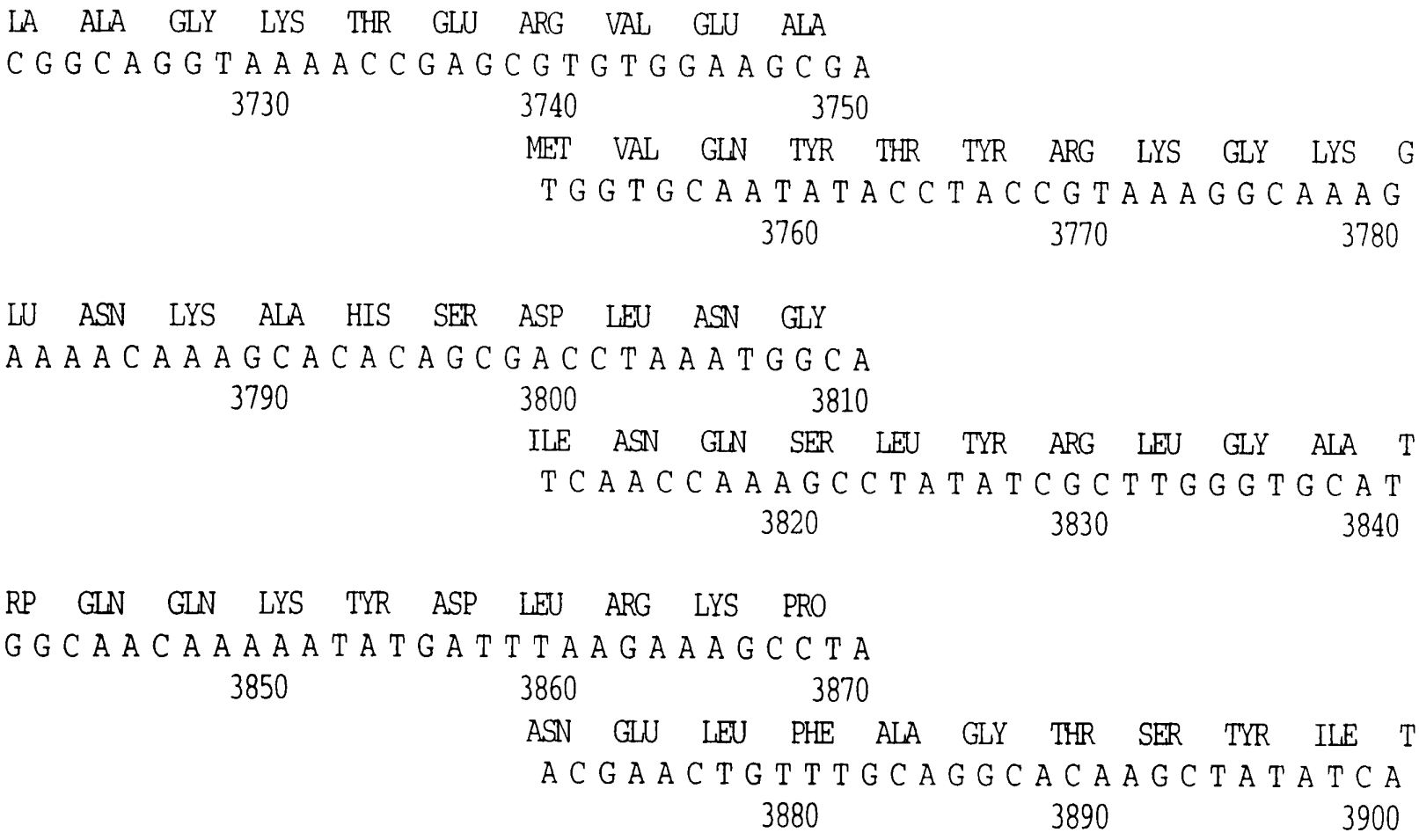


FIG.4V

```
HR  GLU  SER  CYS  LEU  ALA  SER  ASP  ASP  PRO
C C G A A A G C T G T T G G C A A G T G A T G A C C C A A
          3910      3920      3930

LYS  SER  CYS  VAL  GLN  TYR  PRO  TYR  VAL  TYR  T
A A G C T G C G T A C A A T A C C C T T A T G T C T A C A
     3940      3950      3960

HR  LYS  ALA  ARG  PRO  ASP  GLY  ILE  GLY  ASN
C C A A A G C C C C G A C C A G A T G G T A T C G G C A A T C
          3970      3980      3990

ARG  ASN  PHE  SER  GLU  LEU  SER  ASP  ALA  GLU  L
G C A A T T T T T C T G A G T T A A G C G A T G C T G A A A
     4000      4010      4020

YS  ALA  GLN  TYR  LEU  ALA  SER  THR  HIS  PRO
A A G C A C A A T A T T T G G C G T C C A C G C A C C C C C
          4030      4040      4050

HIS  GLU  VAL  VAL  SER  ALA  LYS  ASP  TYR  THR  G
A T G A G G T T G T C T C T G C C A A A G A T T A T A C A G
          4060      4070      4080
```

FIG.4W

```
LY   THR   TYR   ARG   LEU   LEU   PRO   ASP   PRO   MET
G C A C T T A T C G G T T G T T A C C T G A C C C C A T G G
             4090                4100                4110
                                  ASP   TYR   ARG   SER   ASP   SER   TYR   LEU   ALA   ARG   L
                                   A C T A T C G T T C A G A C T C G T A T T T G G C A C G C C
                                               4120                4130                4140

EU   ASN   ILE   LYS   ILE   THR   PRO   ASN   LEU   VAL
T T A A C A T C A A A A T C A C C C C A A A T T T G G T C A
             4150                4160                4170
                                  SER   LYS   LEU   LEU   LEU   GLU   ASP   THR   LYS   GLN   T
                                   G T A A A C T G T T A T T A G A A G A C A C C A A G C A A A
                                               4180                4190                4200

HR   TYR   ASN   ILE   ARG   ASP   MET   ARG   HIS   CYS   S
C A T A C A A C A T T C G T G A T A T G C G T C A T T G T A
             4210                4220                4230
                                  ER   TYR   HIS   GLY   ALA   ARG   LEU   GLY   ASN   ASP   G
                                   G T T A T C A T G G G G C A A G A T T G G G C A A T G A C G
                                               4240                4250                4260
```

FIG.4X

```
LY  LYS  PRO  ALA  ASN  GLY  GLY  SER  ILE  VAL
G T A A G C C C T G C C C A A T G G G C G G G C T C C C A T T G T C C C
                4270                 4280                 4290

                LEU  CYS  ASP  ASP  TYR  GLN  GLU  TYR  LEU  ASN  A
                T T T G C G A T G A T T A T C A A G A G T A T C T T A A A T G
                4300                 4310                 4320

LA  ASN  ASP  ALA  SER  GLN  ALA  SER  PHE  ARG
C C A A T G A C G C C A T C A C A A G C C A T C A T T T A G A C
                4330                 4340                 4350

                PRO  GLY  ALA  ASN  ASP  ALA  PRO  ILE  PRO  LYS  L
                C A G G G G C T A A T G A C G C C C C C A T T C C A A A A C
                4360                 4370                 4380

EU  ALA  TYR  ALA  ARG  SER  SER  VAL  PHE  ASN
T G G C T T A T G C C A G A A G C A G T G T G T T T A A C C
                4390                 4400                 4410

                GLN  GLU  HIS  GLY  LYS  THR  ARG  TYR  GLY  LEU  G
                A A G A G C A T G G C A A A A C T C G C T A T G G G T T A G
                4420                 4430                 4440
```

FIG.4Y

```
LY  PHE  GLU  PHE  LYS  PRO  ASP  THR  PRO  TRP
G T T T G A G T T T A A G C C C T G A C A C G C C C A T G G G T
                4450                 4460                 4470

PHE  LYS  GLN  ALA  LYS  LEU  ASN  LEU  HIS  GLN  G
T T A A A C A A G C A A A A T T A A A C C T A C A T C A A C
          4480                 4490                 4500

LN  ASN  ILE  GLN  ILE  ILE  ASN  HIS  ASP  ILE
A A A A T A T C C A A A T C A T T A A C C A T G A C A T T A
          4510                 4520                 4530

LYS  LYS  SER  CYS  SER  GLN  TYR  PRO  LYS  VAL  A
A A A A A T C G T G C A G C C C A A T A T C C C A A G G T G G
          4540                 4550                 4560

SP  LEU  ASN  CYS  GLY  ILE  SER  GLU  ILE  GLY
A T T T A A A T T G T G G C A T C A G T G A A A T T G G G C
          4570                 4580                 4590

HIS  TYR  GLU  TYR  GLN  ASN  ASN  TYR  ARG  TYR  L
A T T A T G A A T A T C A A A A C A A T T A C C G T T A T A
          4600                 4610                 4620
```

FIG.4Z

```
YS  GLU  GLY  ARG  THR  SER  LEU  THR  GLY  LYS
A A G A A G G G C G T A C C A G T T G A C A G G C A A A C
          4630      4640      4650      4660

LEU  ASP  PHE  ASN  PHE  ASP  LEU  LEU  GLY  GLN  H
T T G A T T T T A A T T T T G A C C T G C T G G G C C A G C
          4670      4680      4690      4700

IS  ASP  LEU  THR  VAL  LEU  ALA  GLY  ALA  ASP
A C G A T T T G A C G G T G T T G G C T G G T G C A G A T A
          4710      4720      4730      4740

LYS  VAL  LYS  SER  GLN  PHE  ARG  ALA  ASN  ASN  P
A A G T T A A A A G C C A A T T T C G T G C C A A C A A C C C
          4750      4760      4770      4780

RO  ARG  ARG  THR  ILE  ILE  ASP  THR  THR  GLN
C C A G A C G C A C A A T C A T T G A C A C C C C A A G
          4790      4800

GLY  ASP  ALA  ILE  ILE  ASP  GLU  SER  THR  LEU  T
G C G A T G C C C A T C A T T G A T G A A A G C A C G C T G A
```

FIG.4A'

```
HR  ALA  GLN  GLU  GLN  ALA  LYS  PHE  LYS  GLN
C A G C A C A G G A G C A A G C C A A A T T T A A G C A A T
          4810              4820              4830
                            SER  GLY  ALA  ALA  TRP  ILE  VAL  LYS  ASN  ARG  L
                             C A G G G G C A G C A T G G A T T G T C A A A A A T C G C T
                                        4840              4850              4860

EU  GLY  ARG  LEU  GLU  GLU  LYS  ASP  ALA  CYS
T A G G A C G C T T A G A A G A A A A A G A C G C C T G T G
          4870              4880              4890
                            GLY  ASN  ALA  ASN  GLU  CYS  GLU  ARG  ALA  PRO  I
                             G C A A T G C C A A T G A A T G T G A A C G C G C G C C C A
                                        4900              4910              4920

LE  HIS  GLY  SER  ASN  GLN  TYR  VAL  GLY  ILE
T T C A T G G C A G T A A C C A A T A T G T G G G C A T T A
          4930              4940              4950
                            ASN  ASN  LEU  TYR  THR  PRO  ASN  ASP  TYR  VAL  A
                             A C A A C C T T T A T A C A C C A A A T G A T T A T G T G G
                                        4960              4970              4980
```

FIG.4B'

SP LEU SER PHE GLY GLY ARG LEU ASP LYS
A T T T A A G T T T T G G T G G A C G C T T G G A T A A A C
4990 5000 5010

GLN ARG ILE HIS SER THR ASP SER ASN ILE I
A A C G C A T T C A C A G C A C C G A T T C A A A C A T C A
5020 5030 5040

LE SER LYS THR TYR THR ASN LYS SER TYR
T C A G C A A A A C T T A C A C C A A C A A A A G C T A T A
5050 5060 5070

ASN PHE GLY ALA ALA VAL HIS LEU THR PRO A
A T T T T G G A G C G G C G G T T C A T C T G A C A C C T G
5080 5090 5100

SP PHE SER LEU LEU TYR LYS THR ALA LYS
A T T T T A G C C T G T T G T A T A A A A C T G C C A A A G
5110 5120 5130

GLY PHE ARG THR PRO SER PHE TYR GLU LEU T
G C T T T C G T A C G C C A A G T T T T T A T G A A C T G T
5140 5150 5160

FIG.4C'

```
YR  ASN TYR ASN SER THR ALA ALA GLN HIS
A C A A C T A T A A C A G C A C C G C C C C A G C A T A
          5170      5180      5190

LYS ASN ASP PRO ASP VAL SER PHE PRO LYS A
A A A A T G A C C C T G A T G T C T T T C C C A A A C
    5200      5210      5220

RG  ALA VAL ASP VAL LYS PRO GLU THR SER
G A G C G G T T G A T G T C A A A C C T G A A A C T T C C C A
          5230      5240      5250

ASN THR ASN GLU TYR GLY PHE ARG TYR GLN H
A T A C C A A T G A A T A C G G C T T T C G C T A T C A G C
    5260      5270      5280

IS  PRO TRP GLY ASP ILE GLU MET SER MET
A C C C T T G G G G G A T A T T G A G A T G A G C A T G T
          5290      5300      5310

PHE LYS SER ARG TYR LYS ASP MET LEU ASP L
T C A A A A G C C G T T A C A A G G A C A T G T T A G A T A
    5320      5330      5340
```

FIG.4D'

```
YS  ALA  ILE  PRO  ASN  LEU  THR  LYS  ALA  GLN
AAGCCATACCGAACCTAACCAAAGCCCAGC
          5350          5360          5370
                        GLN  GLU  TYR  CYS  LYS  ALA  HIS  LEU  ASP  SER  A
                         AAGAGTATTGTAAGGCTCATTTGGATTCCA
                                   5380          5390          5400

SN  GLU  CYS  VAL  GLY  ASN  PRO  PRO  THR  PRO
ATGAATGTGTTGGTAATCCACCCACGCCCA
          5410          5420          5430
                        LYS  THR  SER  ASP  GLU  VAL  PHE  ALA  ASN  LEU  T
                         AAACCAGTGATGAGGTATTTGCCAACTTAT
                                   5440          5450          5460

YR  ASN  ALA  THR  ILE  LYS  GLY  VAL  SER  VAL
ATAATGCCACCATCAAAGGGGTGAGTGTCA
          5470          5480          5490
                        LYS  GLY  LYS  LEU  ASP  LEU  HIS  ALA  MET  THR  S
                         AAGGCAAACTGGATTTGCATGCCATGACAT
                                   5500          5510          5520
```

FIG.4E'

```
ER  LYS  LEU  PRO  ASP  GLY  LEU  GLU  MET  THR
C A A A A C T G C C A G A T G G T C T T G A A A T G A C C T
          5530                5540                5550

LEU  GLY  TYR  GLY  HIS  THR  LYS  LEU  GLY  LYS  P
  T G G G T T A T G G T C A T A C C A A A T T G G G G A A A T
          5560                5570                5580

HE  ASP  TYR  ILE  ALA  PRO  LYS  ASP  ALA  ASP
T T G A T T A C A T T G C A C C C A A A G A T G C C G A T G
          5590                5600                5610

GLY  TRP  TYR  GLN  ALA  ARG  PRO  ALA  PHE  TRP  A
  G T T G G T A T C A G G C T C G C C C C T G C T T T T G G G
          5620                5630                5640

SP  ALA  ILE  THR  PRO  ALA  ARG  TYR  VAL  VAL
A T G C C A T C A C C C C A G C G C T A T G G T C G
          5650                5660                5670

GLY  LEU  ASN  TYR  ASP  HIS  PRO  SER  GLN  VAL  T
  G T C T A A A C T A T G A C C C C C A G T C A A G T A T
          5680                5690                5700
```

FIG.4F'

```
RP   GLY  ILE  GLY  THR  THR  LEU  THR  HIS  SER
GGGGCATTGGCACAACTTTAACGCACAGCA
          5710      5720      5730
                    LYS  GLN  LYS  ASP  GLU  ASN  GLU  LEU  SER  ALA  L
                    AACAAAAAAGATGAAAATGAGCTAAGTGCCC
                              5740      5750      5760

EU   ARG  ILE  ARG  ASN  GLY  LYS  ARG  GLU  ILE
TTAGAATCCGAAATGGCAAAAGAGAAATAC
          5770      5780      5790
                    GLN  THR  LEU  THR  HIS  THR  ILE  PRO  LYS  ALA  T
                    AAACCTTAACGCACACAATACCCAAAGCCT
                              5800      5810      5820

YR   THR  LEU  LEU  ASP  MET  THR  GLY  TYR  TYR
ATACCTTACTGGACATGACAGGCTATTATA
          5830      5840      5850
                    SER  PRO  THR  GLU  SER  ILE  THR  ALA  ARG  LEU  G
                    GCCCAACTGAGAGCATCACCGCTCGTCTTG
                              5860      5870      5880
```

FIG.4G'

```
LY   ILE   ASN   ASN   VAL   LEU   ASN   THR   ARG   TYR
G T A T C A A C A A T G T A T T A A A C A C C C G C T A
             5890                5900
                                    THR   THR   TRP   GLU   ALA   ALA   ARG   GLN   LEU   PRO   S
                                  C A C C A C A T G G G A A G C G G C A C G C C A A C T G C C C A
                                 5910                5920                5930                5940

ER   GLU   ALA   ALA   SER   SER   THR   GLN   SER   THR
G C G A A G C T G C A A G C A G T A C C C A A T C A A C C C
             5950                5960                5970
                                    ARG   TYR   ILE   ALA   PRO   GLY   ARG   SER   TYR   PHE   A
                                  G T T A C A T T G C A C C A G G T C G C A G T T A C T T T G
                                                      5980                5990                6000

                                                        ORF 3
LA   SER   LEU   GLU   MET   LYS   PHE   ***     MET   THR
C C A G T C T T G A A A T G A A G T T T T A A T A T G A C C
             6010                6020                6030
                                    CYS   LEU   PRO   LYS   THR   ASN   PRO   ALA   LEU   LYS
                                  T G T T T A C C A A A G A C C A A C C C T G C T T T A A A A
                                                      6040                6050                6060
```

FIG.4H'

```
VAL  LYS  HIS  ARG  PHE  LEU  LYS  GLN  VAL  LEU
G T C A A G C A C A G A T T T T A A A G C A G G T G C T G
                6070                 6080                 6090

LEU  LEU  LEU  CYS  VAL  ASP  THR  LEU  THR  ALA
T T A T T G C T T T G T T G A T A C A T T A A C A G C A
                6100                 6110                 6120

GLN  ALA  TYR  ALA  HIS  SER  HIS  HIS  THR  PRO
C A G G C G T A C G C C C A C A G C C A T C A T A C G C C C
                6130                 6140                 6150

ILE  HIS  THR  PRO  THR  HIS  GLU  LEU  SER  SER
A T T C A T A C A C C C A C G C A T G A G C T G T C A T C T
                6160                 6170                 6180

ALA  ASP  ALA  LEU  SER  ASP  GLU  GLY  LEU  GLY
G C T G A T G C T T T A T C A G A T G A A G G C T T G G G T
                6190                 6200                 6210

LYS  ASP  LEU  GLY  SER  LEU  ASP  SER  PRO  ASP
A A G G A T T T T G G G C A G T T T G G A C A G C C C A G A T
                6220                 6230                 6240
```

FIG.4I'

```
 GLY   LEU   GLY   ASP   GLY   LEU   GLY   ASP   GLY   LEU
G G T T T G G G T G A T G G T T T A G G C G A T G G T T T G
              6250                6260                6270
                                    GLY   ASP   GLY   LEU   LYS   SER   ASP   LYS   THR   PRO
                                   G G T G A T G G C T T A A A A A G T G A T A A A A C C C C T
                                                       6280                6290                6300

 LEU   PRO   ILE   ASN   ALA   LEU   THR   VAL   ASN   GLN
T T A C C C A T C A A C G C C T T G A C C G T T A A T C A G
              6310                6320                6330
                                    SER   ASN   GLU   SER   GLN   PRO   ALA   PRO   PRO   SER
                                   A G C A A T G A G A G C C A G C C T G C C C C A C C G A G C
                                                       6340                6350                6360

 VAL   ASP   VAL   ASN   PHE   LEU   LEU   ALA   GLN   PRO
G T A G A T G T C A A T T T T T T A C T T G C C C A G C C A
              6370                6380                6390
                                    GLU   ALA   PHE   TYR   HIS   VAL   PHE   HIS   GLN   ALA
                                   G A G G C A T T T T A T C A T G T C T T T C A T C A A G C G
                                                       6400                6410                6420
```

FIG.4J'

```
 ILE  VAL  GLN  ASP  ASP  VAL  ALA  THR  LEU  ARG
A T T G T G C A A G A T G A T G T G G C A A C A T T A C G C
          6430           6440           6450
                                LEU  LEU  LEU  PRO  PHE  TYR  ASP  ARG  LEU  PRO
                               T T G T T A T T G C C A T T T T A T G A C C G C C T G C C T
                                                6460           6470           6480

 ASP  ASP  TYR  GLN  ASP  ASP  VAL  LEU  LEU  LEU
G A T G A T T A T C A A G A T G A T G T T T T G T T G T T A
          6490           6500           6510
                                PHE  ALA  GLN  SER  LYS  LEU  ALA  LEU  SER  ASP
                               T T T G C C C A A A G T A A A C T T G C C C T A A G T G A T
                                                6520           6530           6540

 GLY  ASN  THR  LYS  LEU  ALA  LEU  ASN  LEU  LEU
G G C A A T A C C A A A T T G G C A T T G A A T C T G C T G
          6550           6560           6570
                                THR  ASP  LEU  SER  ASN  LYS  GLU  PRO  THR  LEU
                               A C C G A T T T G A G T A A C A A A G A G C C A A C A C T T
                                                6580           6590           6600
```

FIG.4K'

```
THR  ALA  VAL  LYS  LEU  GLN  LEU  ALA  SER  LEU
A C G G C G G T A A A T T A C A A C T T G C T T C C T T G
                6610                 6620                 6630

LEU  LEU  THR  ASN  LYS  HIS  ASP  LYS  HIS  ALA
T T G C T G A C C A A C A A G C A C G A T A A A C A C G C C
                6640                 6650                 6660

GLN  MET  VAL  LEU  ASP  GLU  LEU  LYS  ASP  ASP
C A A A T G G T G C T A G A T G A A C T C A A A G A T G A T
                6670                 6680                 6690

ALA  HIS  PHE  LEU  LYS  LEU  SER  LYS  LYS  GLU
G C C C A C T T T T A A A A T T A A G C A A A A A A G A G
                6700                 6710                 6720

GLN  ARG  TRP  VAL  LEU  SER  GLN  SER  ARG  TYR
C A A A G A T G G G T G C T A T C G C A A A G T C G C T A T
                6730                 6740                 6750

LEU  HIS  LYS  LYS  TYR  LYS  MET  GLY  LEU  ASP
T T A C A T A A A A A A T A T A A A A T G G G C T T G G A T
                6760                 6770                 6780
```

FIG.4L'

```
LEU  GLY  ILE  ASN  TYR  LEU  HIS  LEU  ASP  ASN
T T G G G C A T C A A C T A T C T G C A T T T G G A T A A T
                6790                 6800                 6810

ILE  ASN  ALA  ALA  SER  THR  ILE  THR  GLN  PRO
A T C A A C G C C G C C T C C A C C A T C A C C C A G C C C
                6820                 6830                 6840

ASN  ILE  LYS  LYS  ASP  ALA  PRO  LYS  PRO  ALA
A A C A T T A A A A A G A T G C C C C A A A A C C T G C T
                6850                 6860                 6870

HIS  GLY  LEU  ALA  LEU  SER  LEU  GLY  VAL  ASN
C A T G G G C T T G C C T T A T C G C T T G G T G A A T
                6880                 6890                 6900

LYS  TYR  THR  PRO  LEU  SER  HIS  GLY  MET  SER
A A A T A C A C G C C G C T T A G T C A T G G C A T G A G T
                6910                 6920                 6930

ILE  TYR  THR  ALA  LEU  ASP  VAL  ASP  GLY  LYS
A T T T A T A C A G C C C T A G A T G T T G A T G G T A A A
                6940                 6950                 6960
```

FIG.4M'

```
 PHE   TYR   ASP   ASP   LYS   SER   HIS   ASN   GLU   LEU
T T T T A T G A T G A C A A A A G C C A C A A T G A A C T G
             6970                6980                6990
                                   ALA   VAL   PHE   ALA   HIS   ALA   GLY   LEU   ARG   LYS
                                  G C G G T T T T T G C T C A T G C T G G A C T A A G A A A A
                                                      7000                7010                7020

 ASP   HIS   GLN   LYS   GLY   TYR   VAL   ASP   VAL   VAL
G A T C A C C A A A A A G G T T A T G T T G A T G T C G T A
             7030                7040                7050
                                   PRO   PHE   VAL   GLY   ARG   ILE   PHE   ALA   THR   ASN
                                  C C T T T T G T T G G G C G T A T T T T T G C C A C C A A T
                                                      7060                7070                7080

 GLN   GLN   HIS   GLY   ARG   LEU   SER   PRO   ARG   LYS
C A G C A G C A T G G C A G A T T A T C C C C C A G A A A A
             7090                7100                7110
                                   ASP   SER   GLN   GLY   VAL   ALA   PHE   GLY   SER   HIS
                                  G A C A G T C A G G G C G T G G C G T T T G G C A G C C A T
                                                      7120                7130                7140
```

FIG.4N'

```
HIS ARG ILE ASN ASP LYS TRP GLN ASN ALA
CATCGGATCAATGATAAATGGCAAAATGCG
      7150      7160      7170

PHE PHE ALA ARG MET GLU LYS GLY ASN TYR
TTTTTGCACGCATGGAAAAGGCAATTAT
      7180      7190      7200

THR GLU HIS TYR GLN GLY TYR ASP GLY LYS
ACCGAGCATTATCAAGGTTATGGCAAG
      7210      7220      7230

ARG TYR HIS VAL ASN ASP THR ILE LEU LEU
CGTTATCATGTGAATGACACCATTTTGTTG
      7240      7250      7260

GLN ASP GLY PRO ASN ARG ARG TYR SER LEU
CAAGATGGCCCCAAATCGTCGTTACTCTTTG
      7270      7280      7290

GLY VAL GLY TYR GLN LEU SER HIS LEU GLN
GGCGTGGGGTATCAGCTTAGCCATCTGCAA
      7300      7310      7320
```

FIG.4O'

```
 ASP  ALA  THR  LYS  SER  SER  HIS  ALA  THR  LYS
G A T G C A A C A A A A A G C A G T C A T G C C A C A A A G
            7330              7340              7350
                                 ILE  HIS  PHE  GLY  VAL  LEU  GLN  ARG  LEU  PRO
                               A T A C A T T T T G G G G T G T T G C A A A G A T T G C C A
                                              7360              7370              7380

 ASN  GLY  LEU  THR  VAL  GLN  GLY  ARG  VAL  SER
A A T G G T C T G A C C G T G C A A G G T A G A G T G A G T
            7390              7400              7410
                                 ALA  GLU  ARG  GLU  ARG  TYR  HIS  GLY  LYS  LEU
                               G C T G A G C G T G A G C G T T A T C A T G G T A A A T T A
                                              7420              7430              7440

 LEU  ARG  LEU  VAL  ASN  PRO  ASP  ASP  VAL  TYR
T T G C G T C T G G T T A A T C C T G A T G A T G T G T A T
            7450              7460              7470
                                 ARG  THR  ASP  LYS  THR  LEU  THR  LEU  GLN  THR
                               C G C A C A G A T A A A A C C C T A A C C C T A C A A A C C
                                              7480              7490              7500
```

FIG.4P'

```
 SER  ILE  TRP  HIS  LYS  ASP  ILE  HIS  TRP  LEU
TCCATTTGGCACAAAGACATTCACTGGCTT
          7510          7520          7530
                                GLY  LEU  THR  PRO  LYS  LEU  THR  TYR  ARG  TYR
                               GGATTAACGCCAAAGCTGACTTATCGTTAC
                                          7540          7550          7560

SER  LYS  ASN  ASN  SER  ASN  LEU  PRO  ALA  LEU
AGTAAAAATAACAGTAACTTACCAGCACTT
          7570          7580          7590
                                TYR  SER  HIS  ASN  LYS  GLN  ASN  PHE  TYR  LEU
                               TATAGCCATAACAAACAAAATTTTTATTTG
                                          7600          7610          7620

 GLU  LEU  GLY  ARG  SER  PHE  ***
GAGCTTGGTCGGTCGTTTTAA
          7630          7640
```

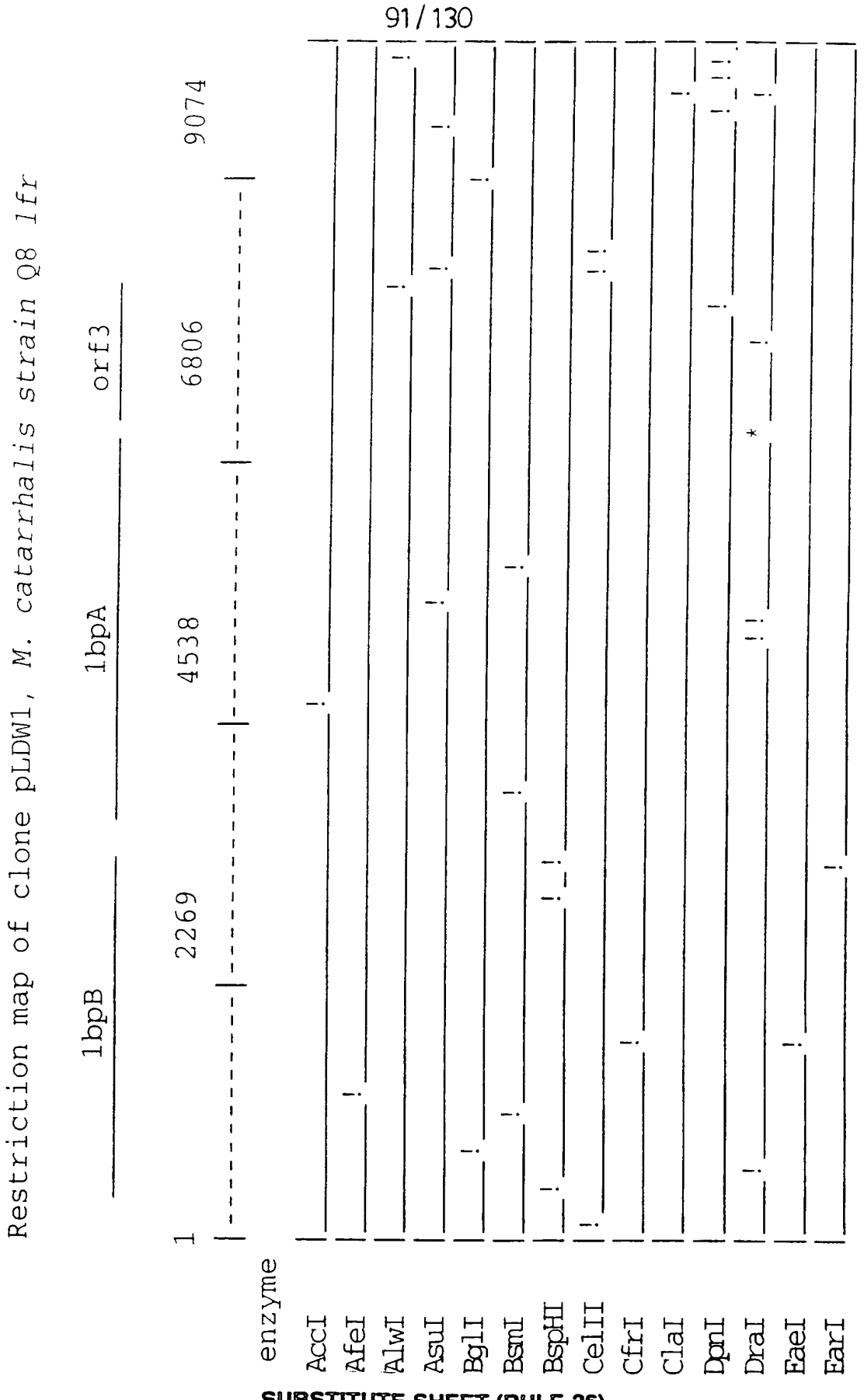


FIG.5A

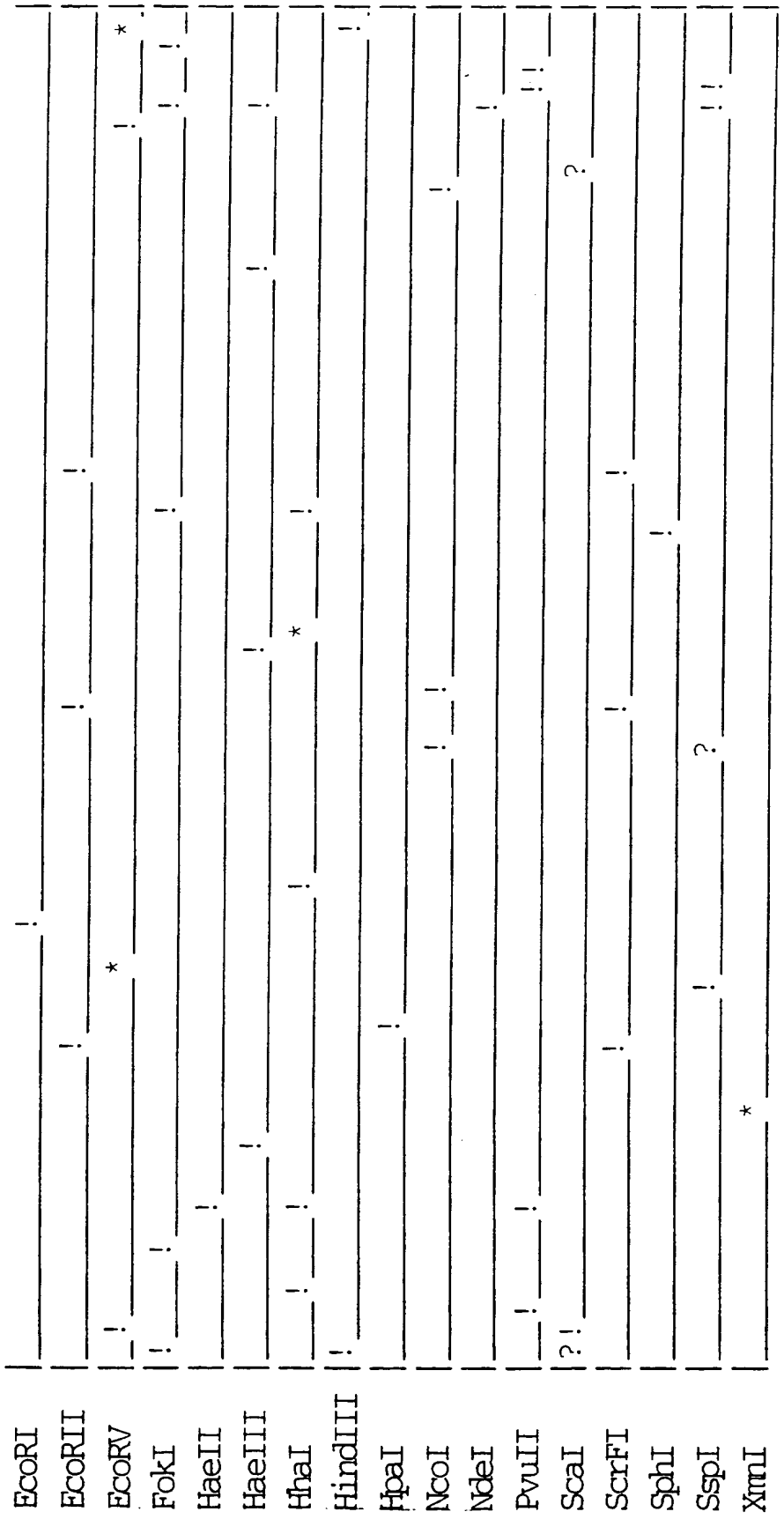


FIG.5B

FIG.6A

Alignment of Lbp2 proteins

```
        10         20         30         40         50         60
MSKSITKTQTPSVHTMTTHRLNLAIKAALF-GVAVLPLSVWAQENTQTDANSDAKDTKTPV
.............................-..............................
.N.KHGFQL.LTA---------..VA..-.PSY.AN.---------E.A.-P..AQ.QS--
.N.KHSFPL.LTA---------...AT.-.PSY.ANS---------E.A.-----Q.QS--
.N.KHGFPL.LTA---------...AT.-.PAY.AQA---------GAA.-L..AQSQS--
                                      70         80         90        100
                                     VYLDAITVTAAPSAPVSRFDTDVTGLGKTVKTADTLAKEQ                4223
                                     ........................................                Q8
                                     --.KEV..R..K---.G.RSKEA.....IA..SE..N...                BNCV
                                     --.KEV..R..K---.G.RSKE......I...SE..N...                H44/76
                                     --.KEV..R..K---.G.RSKEA.....I...SE..N...                FA19

      110        120        130        140        150        160
VQGIRDLVRYETGVSVVEQGRGGSSGFAIHGVDKNRVGITVDGIAQIQSYKDESTKRAGA-
.........................................................-
.L.....T..DP..A.....N.A.G.YS.R.......AVS...V....AFTVQGSLSGYGG
.L.....T..DP..A.....N.A.G.YS.R.......AVS...V....AFTVQGSLSGYGG
.L.....T..DP..A.....N.A.G.YS.R.......AVS...V....AFTVQGSLSGYGG
                                          170        180        190        200
                                     --GSGAMNEIEIENIAAVAINKGGNALEAGSGALGGSVAFHT              4223
                                     --.....................................                  Q8
                                     RG....I....Y...ST.E.D..AGSSDH.......A...R.              BNCV
                                     RG....I....Y...ST.E.D..AGSSDH.......A...R.              H44/76
                                     RG....I....Y...ST.E.D..AGSSDH.......A...R.              FA19
```

93/130

FIG.6B

```
      210        220        230        240        250        260
KDVSDVLKSGKNLGAQSKITYNSKNDHFSQTLAAAGKTERVEAMVQYTYRKGKENKAHSDL
..........N.....................................................
.EAA.LISD..SW.I.A..A.G...RQ.MKS.G.GFSKDGW.GLLIR.E.Q.R.THP.G.I
.EAA.LISD..SW.I.A..A.G...RQ.MKS.G.GFSKDGW.GLLIR.E.Q.R.TRP.G.I
.EAA.LISD..SW.I.A..A.G...RQ.MKS.G.GFSKDGW.GLLIR.E.Q.R.TRP.G.I
                                    270        280        290        300
                                -NGINQSLYRLGAWQQKYDL-RKPNELFAGTSYITESCLAS      4223
                                -.................-.................      Q8
                                AD.VAYGIN..D.FR.T.GI-K..S.GGEYFLAEG..E.KP      BNCV
                                AD.VAYGID..D.FR.T..I-Q.QNKKAEYFLAEG..E.KP      H44/76
                                AD.VAYGID..D.FR.T..IK..TT.P.--FLAEG.NT.KP      FA19
      310        320        330        340        350        360
DDPKSCVQYPYVYTKARPDGIGNRNFSELSDAEKAQYLASTHPHEVVSAKDYTGIYRLLPD
..................................................T.....
VAKVAGNGNYLNNQLN.WVKERIEQNQP..AE.E.MVREAQAR..NL..QA...GG.I...
AAKLAGNGNYLNNQLN.WVEERKKNNQS..AE.E.MVREAQAR..NL..QA...GG.I...
VAKLAGYGIYLNRQLN.WVKERIEQNQP..AE.E..VREAQAR..NL..QA...GG.I...
                                    370        380        390        400
                                PMDYRSDSYLARLNIKITPNLVSKLLLEDTKQTYNIRDM      4223
                                ........................................      Q8
                                ......G.W..K.GYRFGGRHYVGGVF.....R.D....      BNCV
                                ......G.W..K.GYRFGGRHYVGGVF.....R.D....      H44/76
                                ......G.W..K.GYRFGGRHYVGGVF.....R.D....      FA19
```

FIG.6C

```
      410       420       430       440       450       460
RHCSYHGARLGNDGKPANGGSIVLCDDYQEYLNANDASQALFRPGANDAPIPKLAYARSSV
.....................................S.....................
TEKQ.Y.TDEAKKFRDKS.—VYDG..FRDG.YFVPNIEE-WKGDQKLIRGIG.K.S.TK-
TEKQ.Y.TDEATKFSDKS.—VYDG..FRDG.YFVPNIEE-WKGDK.LVKGIG.K.S.TK-
TEKQ.Y.TDEAEKFRDKS.—VYDG..FRDG.YFVPNIEE-WKGDK.LVKGIG.K.S.TK-
                              470       480       490       500
                    FNQEHGKTRY-GLSFEF--KPDTPWFKQAKLNLHQQNIQIIN
                    .....-..G...---.....................
                    .ID..HRR.RM..LYRYENE.YSDN.ADK.V.SFDK.GVATD.
                    .ID..HRR.RM..LYRYENEAYSDN.ADK.V.SFDK.GVATD.
                    .ID..HRR.RM..LYRYENE.YSDN.ADK.V.SFDK.GVATD.

      510       520       530       540       550       560
HDIKKSCSQYPKVDLNCGISEIGHYEYQ—NNYRYKEGRASLTGKLDFNFDL-LGQHDLITVLAG
.....--.....T.....-.....
NTL.LN.AV..A..KS.RA.ADKP.S.DSSDRFH.R.QHNV.NASFEKSLKNKWTK.H..LGF.
NTL.LN.AV..A..KA.RA.ADKP.S.DSSDRFH.R.QHNV.NALFEKSLKNKWTK.H..LGF.
NTL.LN.AV..A..KA.RA.ADKP.S.DSSDRFH.R.QHNV.NASFEKSLKNKWTK.H..LGF.
                              570       580       590       600
                    ADKVKSQFRANNPRRTIIDITTQGDAIIDESTLTAQEQAK
                    .....
                    Y.ASNAIS.PEQLSHNAARISEYSDYT.KGD-------
                    Y.AS.AIS.PEQLSHNAARISEFSDYA.DGKY------
                    Y.AS.AIS.PEQLSHNAARISE-STGF..KNQD-----
```

```
4223
Q8
ENCV
H44/76
FA19
```

```
4223
Q8
ENCV
H44/76
FA19
```

FIG.6D

```
       610        620        630                640        650
FKQSGAAWIVKNRLGRLEEKDA--CGNANECERAP-----IHGSNQYVGINNLYTPNDYVD
.......................--..........-----.....................
----------.YL..KP.VVEGSV..YIETLRSRKCVPRK.N...IHISL.DRFSIGK.F.
----------.YL..KP.VVEGSV..YIETLRSRKCVPRK.N...IHISL.DRFSIGK.F.
----------.Y...KP.VVEGSV..YIETLRSRKCVPRK.N...IHISL.DRFSIGK.F.
                                    660        670        680        690        700
                                   LSFGGRLDKQRIHSTDSNIISKTYTNKSYNFGAAVHLTPDFSLLYK
                                   ...........................................
                                   F.L...Y.RKNFTTSEELVR.GR.VDR.W.S.IVFKPNRH...S.R.
                                   F.L...Y.RQNFTTSEELVR.GR..DR.W.S.IVFKPSRHL..S.R.
                                   F.L...Y.RKNFTTSEELVR.GR.ADR.W.S.IVFKPNRH..VS.R.

      710        720        730        740        750        760
TAKGFRTPSFYELYNYNSTAAQHKNDPDVSFPKRAVDVKPETSNTNEYGFRYQHPWGDVEM
.....................................................I..
ASS.......Q..FGIDIYH-------.YPKGWQRPAL.S.KAANR.I.LQWKGDF.FL.I
ASS.......Q..FGIDIYH-------.YPKGWQRPAL.S.KAANR.I.LQWKGDF.FL.I
ASS.......Q..FGIDIYH-------.YPKGWQRPAL.S.KAANR.I.LQWKGDF.FL.I
                                        770        780        790        800
                                   SMFKSRYKDMLDKAIPNLTKAQ-QEYCKAHLDSNECVGNP
                                   .....................-.................
                                   .S.RN..T..IAV.DHK-..LPN.AGQLTEI.IRDYY---
                                   .S.RN..T..IAV.DHK-..LPN.AGRLTEI.IRDYY---
                                   .S.RN..T..IAV.DQK-..LPDSAGRLTEI.IRDYY---
```

4223
Q8
BNCV
H44/76
FA19

4223
Q8
BNCV
H44/76
FA19

FIG.6E

```
      810        820        830        840        850        860
PTPKTSDEVFANLYNATIKGVSVKGKLDLHAMTSKLPDGLEMTLGYGHTKLGKFDYIAPKD
............................................................
----------.AQ.MSLQ..NIL..I.WNGVYG...E..YT..A.NRI.--------..S
----------.AQ.MSLQ..NIL..I.WNGVYG...E..YT..A.NRI.--------..S
----------.AQ.MSLQ.INIL..I.WNGVYG...E..YT..A.NRI.--------..S
                                    870        880        890        900
                                   ADGWYQA--RPAFWDAITPARYVVGLNYDHPSQVWGIGTTL        4223
                                   .......--..............................        Q8
                                   VSNRPGLSL.SYAL..VQ.S...L.FG..Q.EGK..ANIM.                BNCV
                                   VSNRPDLSL.SYAL..GQ.S...L.FG..Q.EGK..ANIM.        H44/76
                                   VSNRPDLSL.SYAL..VQ.S...L.FG..Q.EGK..ANIM.        FA19

      910        920        930        940        950        960
THSKQKDENELSALRIRNGKRETQTLTHTIPKAYTLLDMTGYYSPTESITARLGINNVLNT
......................I.....................................
.Y..G.NPD..-.YLAGDQ..YS---.KRASSSWSTA.VSA.LNLKKRL.L.AA.Y.IG.Y
.Y..G.NPD..-.YLAGDQ..YS---.KRASSSWSTA.VSA.LNLKKRL.L.AA.Y.IG.Y
.Y..G.NPD..-.YLAGDQ..YS---AGRVTSSWKTA.VSA.LNLKKRL.L.AA.Y.IG.Y
                                    970        980            990       1000
                                   RYTIWEAARQLPSEAASSTQST-----RYIAPGRSYFASLEMKF*                4223
                                   ......................-----.................*                Q8
                                   ..V...SL..T-A.STANRHGGDSNYG..A....NFSLA.....*                BNCV
                                   ..V...SL..T-A.STANRHGGDSNYG..A....NFSLA.....*                H44/76
                                   ..V...SL..T-A.STANRHGGDSNYG..A....NFSLA.....*                FA19
```

FIG.7A

Alignment of M. catarrhalis Lbp2 proteins

```
       10        20        30 ↓      40        50        60
MSTVKTPHIFYQKRTLSLAIASIFAALVMTGCRSDDISVNAPNVTQLPQGTVSPTPNTGH
.....P.....................................................I.....
.....V.....................V.....................I.....

                                          70        80        90        100
                                DNANNTNNQGNNTDNSTSTTDPNGDNNQLTQAQKTAAAAG          Q8
                                ..T.....................................          4223
                                ..T.....................................          VH19

      110       120       130       140       150
FFVMGKIRDTSEKNDPDYSDDLKQQW----LGKLYVGIDAHRPDGIGKGKNLRQPITAND
..........P.......N..V...----Q.............T...........
.................TK..QGSVHTAGQ.LQ.L.TKEP....T.T.........D.
                                 160       170       180       190       200
                                IKPLYFNKFPALSDLHLDSERHRFDPQKINTIKVYGYGNLTTPS      Q8
                                ........................K.L.............          4223
                                .T....D...KI.....ENSE.V..AK.A.N..I....A.SS.A      VH19

      210       220       230       240       250       260
NNNTHINHQQADNKKNNKPVDPYENIRFGYLELQGSSLTQKNADNQNEQDR-IPKPMPILF
K...Y.....................................P.DK..-.........
K.P.YM.Y..EQ.I..K..G.D.Q......M..RELD.NK.G....SDKN.A.IFTT.T..
                                          270       280       290       300
                                YHGENASSQLPSAGKFNYTGNWLYLSDVKKRPALSASDER          Q8
                                .....................................D.          4223
                                .......TH..K....D.E......T......F.DKT.DK         VH19
```

FIG.7B

```
                310       320       330       340       350       360
VGVYLNASGKANEGDVVSAAHIYLNGFQYKHTPATYQVDFDTNSLTGKLSYYIDNPNQQNN
          .....S.....................................................TA
          ..T.F.STR.S....L.....S.K.....S....Q.T.K.....K.TA
                                    370       380       390       400
                                    KGEYLKSQFDITKKVNETDVYQIDAKINGNRFVGTAKSLV
                                    Q.K.I.....................................
                                    D.R.IR.....D.....A...E.....T.....I

                410       420       430       440       450       460
NEKTQTAPFIKELFSKKANPNNPNPNSDTLEGGFYGESGDELAGKFLSNDNASYVVFGGK
  .N.E.....................................................
DDN.N....V.....................D.....................TF.....
                                    470       480       490       500
                                    RDKTTKPVATKTVYFSAGFEKPSTSFVDNETIGGIIDRKG
                                    ....D.....................R..NS.K
                                    .....E.....T.....G..E..S...G..

          510       520       530       540       550
LN---NHINEDEIIP-SDDSYYGYTWGKPEKQFTKKVSSSTQVVPAYFGQHDKFYFNGN
..DAVNEK.DNGD.PT-..ER.DEFP..EKKAE.....A.....
..DEVN.Q.-....TV.V.NKE..E.NY.R.N.....INA.V.KN.....
                                    560       570       580       590       600
                                    YYDLSASRVDKLAPADAVKANQSIKEKYPNATLNKDTNQVTALVLQ
                                    .....S.....
                                    .....KEAN..GVSQDTST.K..LA...D.KVST..K..K....
```

Q8
4223
VH19

Q8
4223
VH19

Q8
4223
VH19

FIG.7C

```
      610       620       630       640       650       660
EAKDNKPYTAIRAKSYQHISFGETLYNDANQTPTRSYFVQGGRADTSTTLPQAGKFTYNG
.....................................................K.........
Q...-.....H....D.....V....NKGN.....Q..V..Q..S.........
                    LK.I.TAEAD-I..T..AR.T.
                      670       680       690       700
                    LWAGYLTQKKDKGYSDNAETIKEKGHPGYLLTENFTPEDD
                    .....I.....N.E....K...QD.....D.....
                    ..........KDED...Q..LKD.I..KD.I.Q..
                    T.EARISKPIQWDDNHADKKAA----------------
      710       720       730       740       750
DDD--LTASDDSQDDNTHGDDDLIASDDSQDDDALGDDDSDDLGDGADDDAAGKVYHA
...---.....DA.....A.....
...DDS.....T.....
-------------------------------------------------------
      760       770       780       790       800
GNIRPEFENKYLPTNEPTHEKTFALDGKNKAKPEVDPNTNSLTG
.....D...D.....
.....D.N.D.....
----------.E.D...GEK.IS.
```

Q8
4223
VH19
H44/76

FIG.7D

```
      810        820        830        840        850
KLNDERGDIV-FDIKNGKIDGTGFTAKADVPNYREEV-GNNQGGGFLYNIKDIDVKG
.........-..........................-...................
.........-..........................-...................

T.TEKN.VQPA.H.E..V.E.N..H.T.RTRDNGINLS..DSTNPPSFKANNLL.T.
                           RDNGINLS..GSTNPPSFKADNLL.T.
                                  860        870        880
                                QFFGTNGEELAGQLHHDKGD--GINDTAEKA----------------  Q8
                                ..............QY....--.........----------------  4223
                                R..................--..T......----------------  VH19
                                G.Y.PKA...G.IIFNNND.KSL..TEGT.NKVE-ADVDVDVDVDVD  BNCV
                                G.Y.PQA...G.IIFNNND.KSL..TEDT.NEAE-AEVENEAGVG--  H44/76
                                G.Y.PQAA..G.IIFNNND.KSL..TEDI.NEVENEADVG-------  FA19

                      890
----------------GAVFGAVKD-----K*                                                 Q8
----------------.........-----.*                                                 4223
----------------.........-----.*                                                 VH19
ADADVEQLKPEVKPQF.V....K..NKEVE.*                                                 BNCV
-----EQLKPEAKPQF.V....K..NKEVE.*                                                 H44/76
-----EQLEPEVKPQF.V....K..NKEVE.*                                                 FA19
```

FIG.8A *M. catarrhalis strain 4223 Tbp2/Lbp2 comparison*

```
     MKHIPLT--TLCVAISA----VLLTACGGSGGS-NPPAPT--------PIPNASGSGNTG    Tbp2
MSTVKTP..FYQKR..SL..ASIFAALVM.G.RSDDI.V.A.NV.QLPQGTVS....TGHD-..N    Lbp2
       **     ** *****    ****** *   * * *  *         ******   **

NTGNAGG-TDNTANAGNTGGTNSGTGSANTPEPKYQDVPTEKNEKDKVSSIQEPAMG-----YGM    Tbp2
..N.Q.NN...STSTTDPN.D.NQLTQ.QKTAAAAGFFVMG.I-R.TSPKNDPDYSNDLVQQWQG    Lbp2
** * *  ***** ***  * *   * *              *  **     *          *

ALSKINLHNRQDTPLDEKNIITLDGKKQVAEGKKSPLPFSLDVENKLLDGYIAKMNVADKNAIGD    Tbp2
K.YVGIDAH.P.GIGTG..LRQPITANDIKPLYFNKF.ALS.L-------------HL.SERHRF    Lbp2
 *       * **    ***    *  **        *   **                 * *

RIKKGNKEISDEELAKQIKEAVRKSHEFQQVLSSLENKIFHSNDGTTKATTRDLKYVDYGYYLAN    Tbp2
DP..L.TIKVYGYGNLTTPSKNNTYINH..ADNKKN..PVDPYENIRFGYLELQGSSLTQKNADT    Lbp2
  ** *                        **     ***     *    *

DGNYLTVKTDKLWNLGPVGGVFYNGTTTAKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENSQ    Tbp2
PNDKDRIPK-------.MPIL..H.ENASSQ..SAGKFN.T.N.LYLS..KK.PALSASDDRV--    Lbp2
               **  *** *  ** ****     * * * *****  *    *

AGWYYGASSKDEYNRLLTKEDSAPDGHSGEYGHSSEFTVNFKEKKLTGKL-------------F    Tbp2
-.V.LN..G.SNEGDVVSAAHIYLN.FQYKHTPAT-YQ.D.DTNS.....SYYDNPNQQTAQGKY    Lbp2
 * *  **** *   ***      **     * ** * ***    *****               *

SNLQDRHKGNVTKTERYDIDANIHGNRFRGSATA-----SNKNDTSKHPFTSDAN--------NR    Tbp2
IKS.FDTTKK.NE.DV.Q...K.N....V.T.KSLVNENTETAPFI.EL.SKK..PNNPNPNSDT    Lbp2
   *      *  ** * *** * **** *** *     **     *  **  **        *
```

SUBSTITUTE SHEET (RULE 26)

FIG.8B

```
LEGGFYGPKGEELAGKFLTNDNKLFGVFGAKRESKAEEKTEAILDAYALGTFNTSNATTFTPFTE    Tbp2
.......ES.D.......S...ASYV...G..DKTDKPVATKTVYFS.AFEKP-----------    Lbp2
*******  **************  * *******       *   *   *

KQLDNFGNAKKLVLGSTVIDLVPTDATKNEFTKDKPESATNEAGETLMVNDEVSVKTYGKNFEYL    Tbp2
------------------------------------------------------------------    Lbp2

KFGELSIGGSHSVFLQGERTATTGEKAVPTTGTAKYLGNWVGYITGKDTGTGTGKSFTDAQDVAD    Tbp2
------------------------------------------------------------------    Lbp2

FDIDFGNKSVSGKLITKGRQDPVFSITGQIAGNGWTGTASTTKADAGGYKIDSSSTGKSIAIKDA    Tbp2
-----------------------------------------.IKNG...GTGFTAKADVPNY    Lbp2
                                          *    ****** *    * *

------NVTGGFYGPNAN-EMGGSFTHN-----ADDS-------------KASVVFGTKRQQEVK*    Tbp2
REEVGN.QG...LYNIKDIDVK.Q.FGTNGEEL.GQLQYDKGDGINDTAE..GA...AVKD---.*    Lbp2
      * ****     *  * * *        * *               *** **** **   *
```

FIG.9A Alignment of *M. catarrhalis Lbp3 sequences*

```
        10         20         30         40         50
MTCLPKTNPALKVKHRFLKQVLLLLCVDTLTAQAYAHSHHTPIHTPTHEL    4223
..................................................     Q8

        60         70         80         90        100
PSADALSDEGLGKDLGSLDSLDSPDGLGDGLGDGLGDGLKSDKAPLPINA    4223
S.....................................T......        Q8

       110        120        130        140        150
LTAHQTNESQPAPPSVDVNFLLAQPEAFYHVFHQAIVQDDVATLRLLLPF    4223
..VN.S..............................................   Q8

       160        170        180        190        200
YDRLPDDYQDDVLLLFAQSKLALSDGNTKLALNLLTDLSNKEPTLTAVKL    4223
..................................................     Q8

       210        220        230        240        250
QLASLLLTNKHDKHAQMVLDELKDDAHFLKLSKKEQRWVLSQSRYLHKKY    4223
..................................................     Q8

       260        270        280        290        300
KMGLDLGINYLHLDNINAASTITQPNIKKDAPKPAHGLALSLGVNKYTPL    4223
..................................................     Q8

       310        320        330        340        350
SHGMSIYTALDVDGKFYDDKSHNELAVFAHAGLRKDHQKGYVDVVPFVGR    4223
..................................................     Q8
```

FIG.9B

```
      360       370       380       390       400
IFATNQQHGRLSPRKDSQGVAFGSHHRINDKWQNAFFARMEKGNYTERYQ  4223
.....................................H..            Q8

      410       420       430       440       450
GYDGKRYHVNDTILLQDGPNRRYSLGVGYQLSHLQDATKSSHATKIHFGV  4223
..................................................  Q8

      460       470       480       490       500
LQRLPNGLTVQGRVSAERERYHGKLLRLVNPDDVYRTDKTLTLQTSIWHK  4223
..................................................  Q8

      510       520       530       540
DIHWLGLTPKLTYRYSKNNSNLPALYSHNKQNFYLELGRSF*          4223
..................................................*  Q8
```

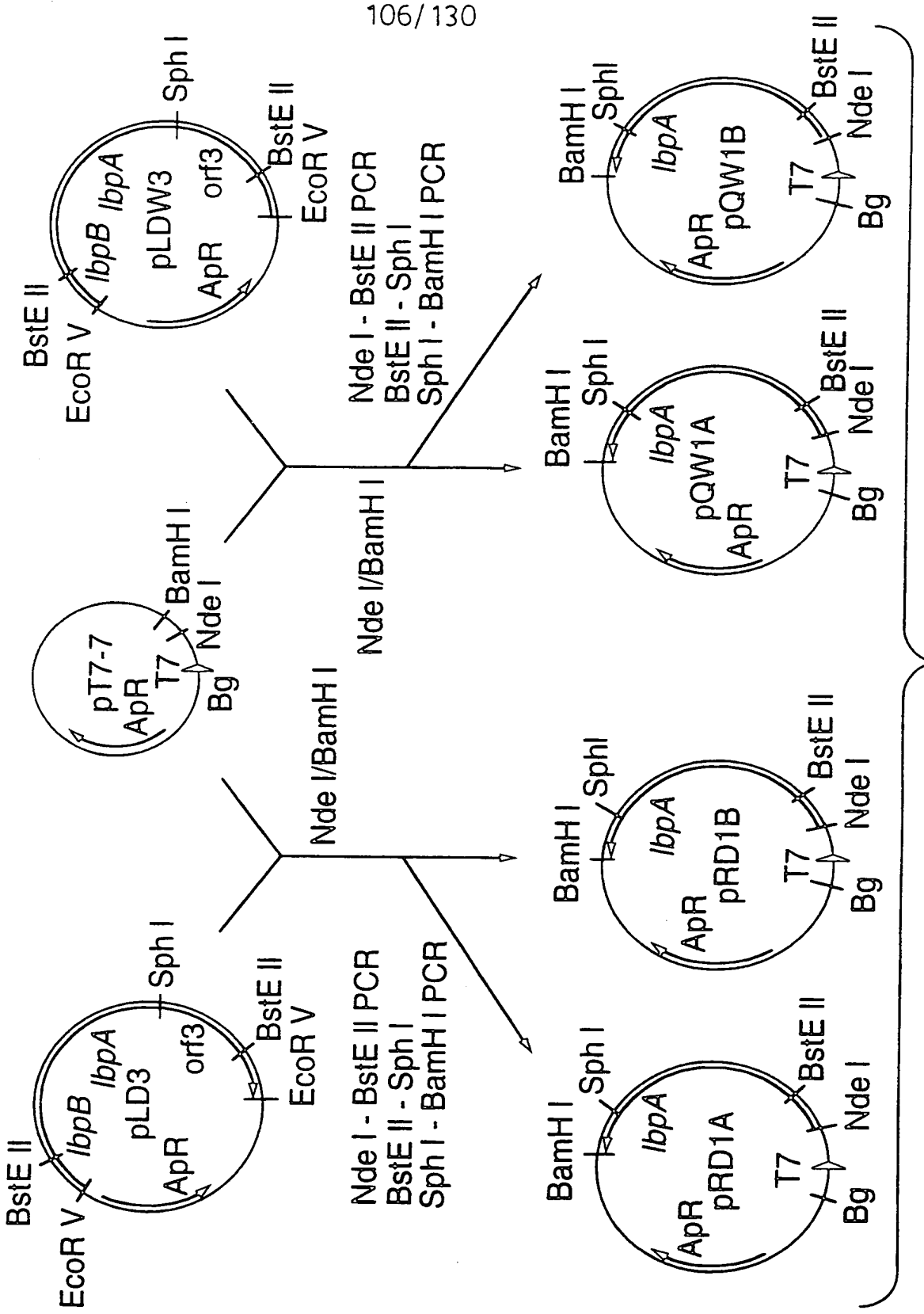


FIG.10

EXPRESSION OF Q8 AND 4223 rLbp 1 PROTEINS.

FIG. 11A.

Q8 rLbp1

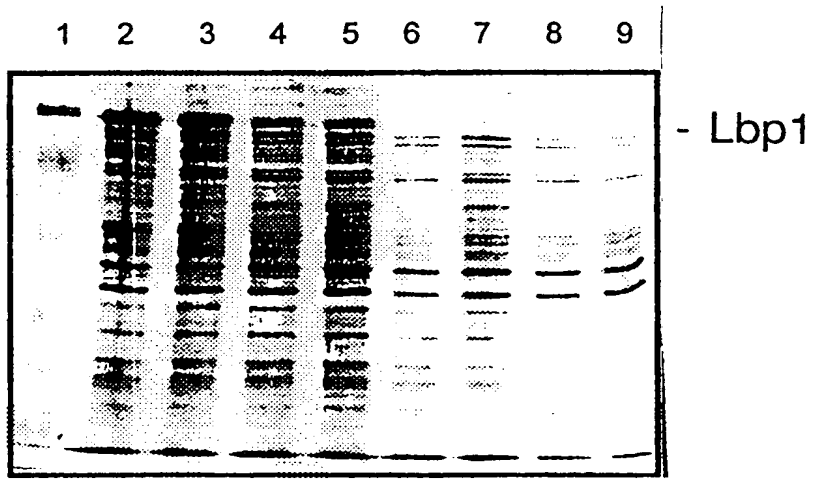
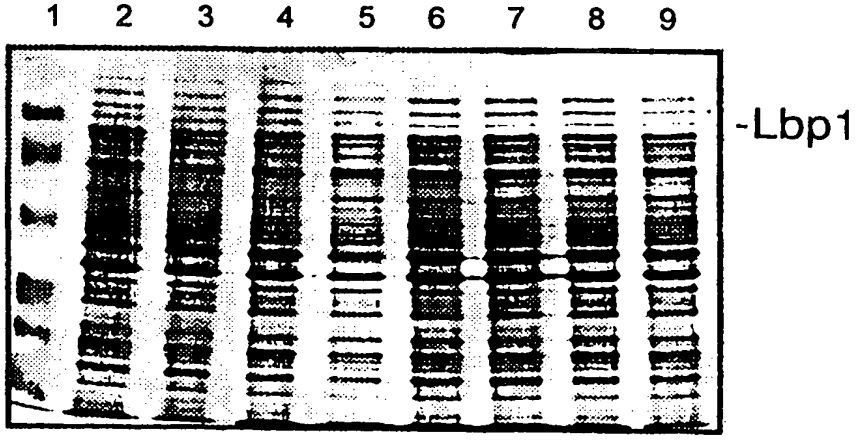


FIG.11B.

rLbp1



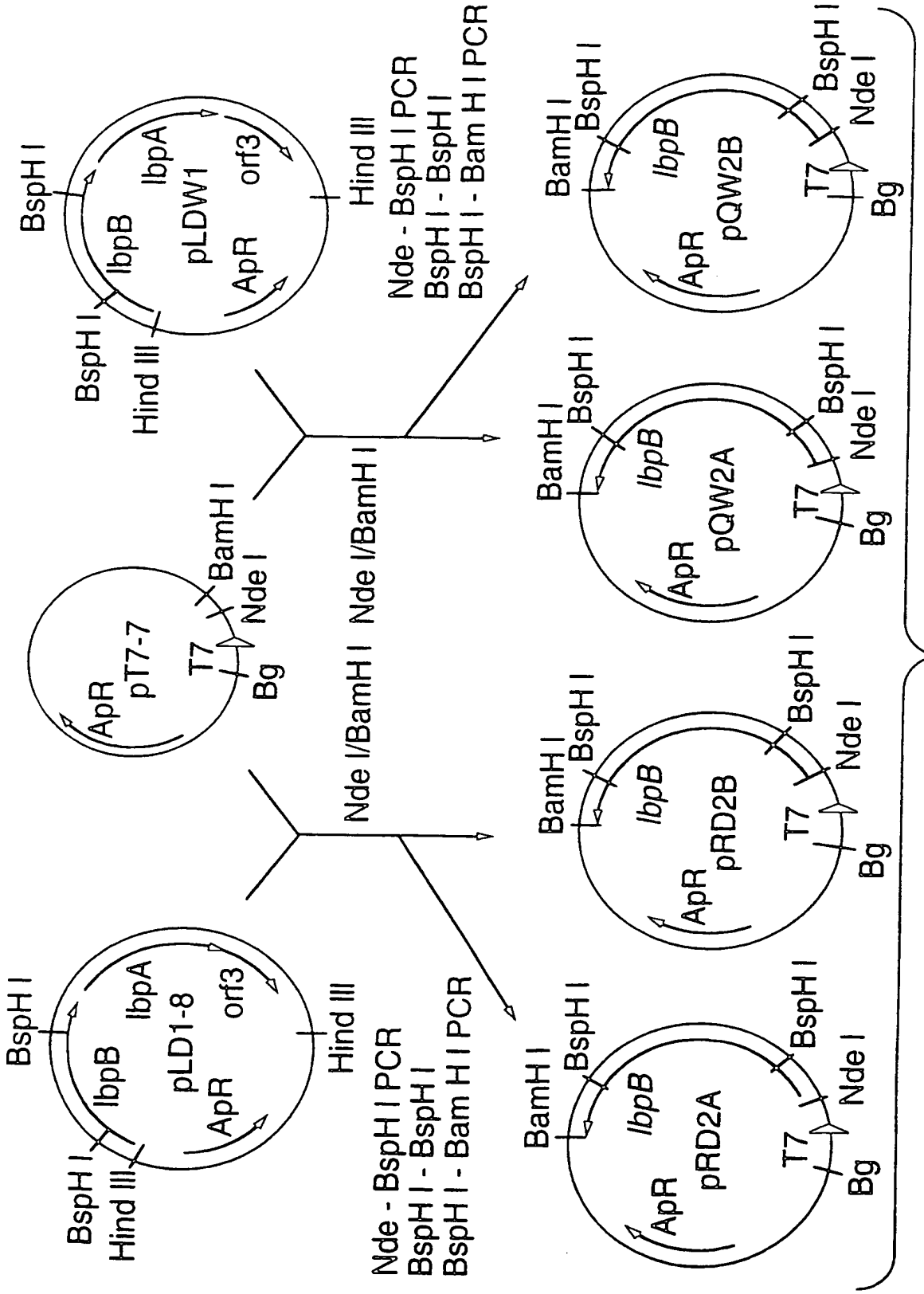


FIG.12

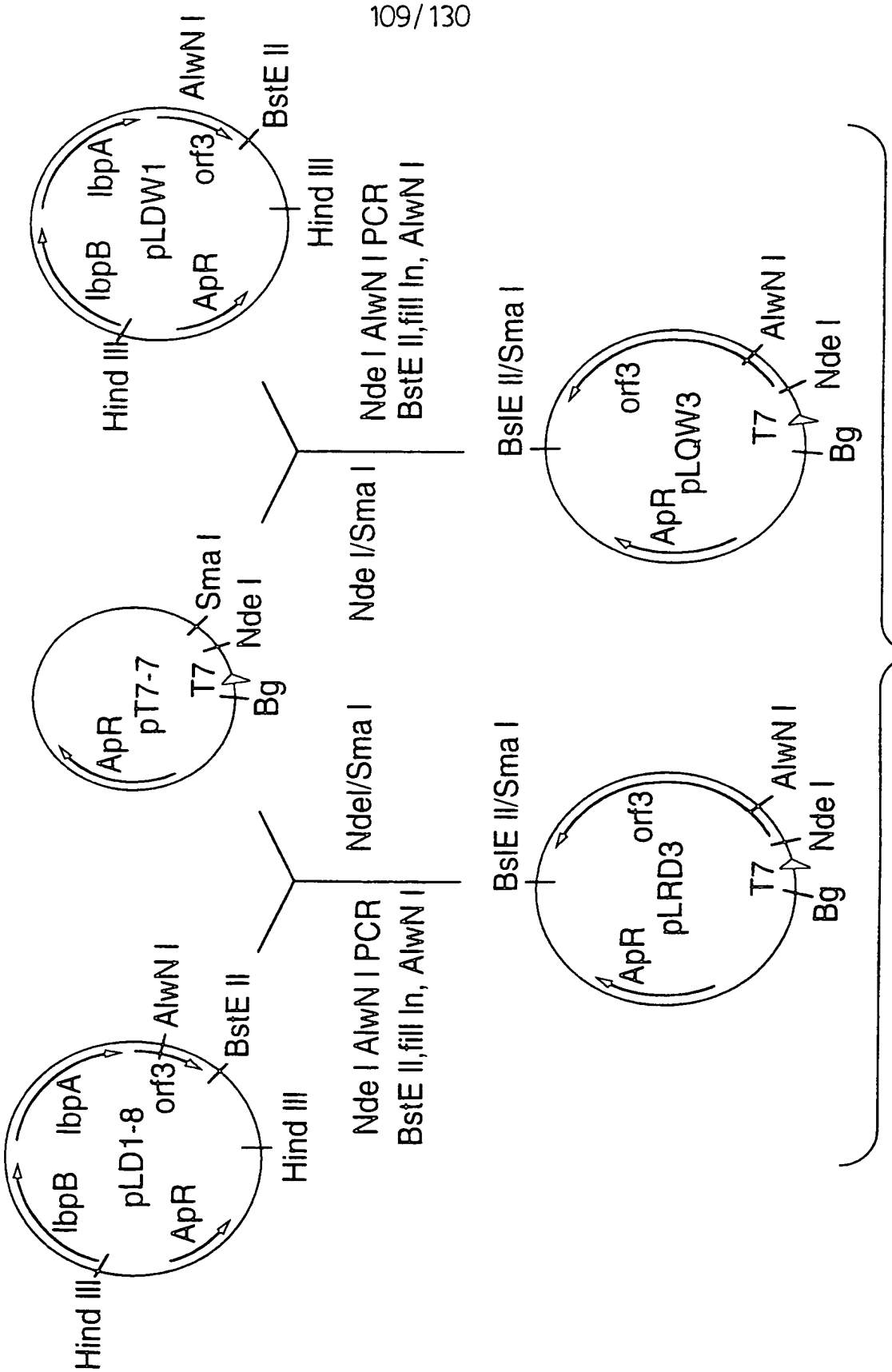
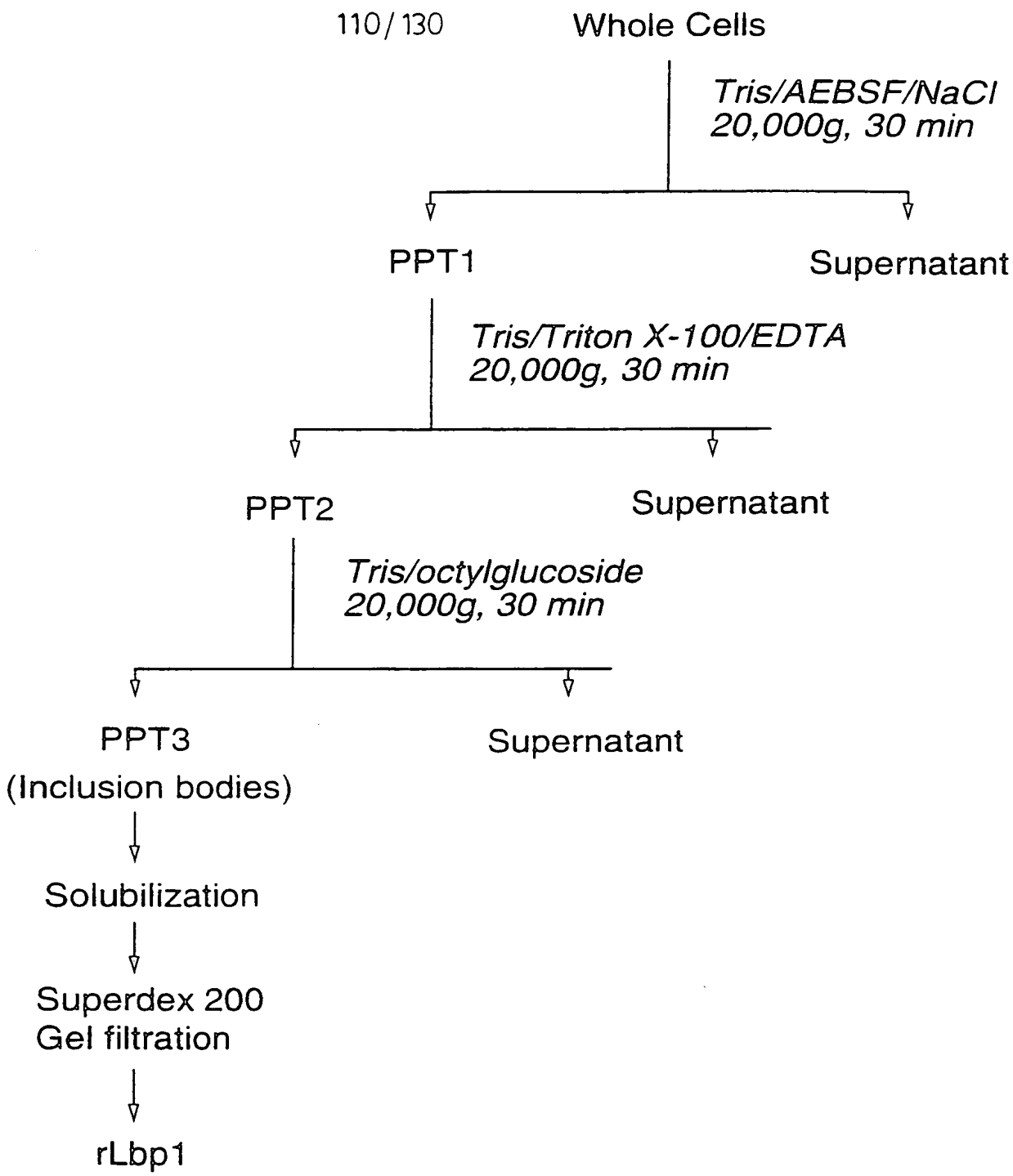


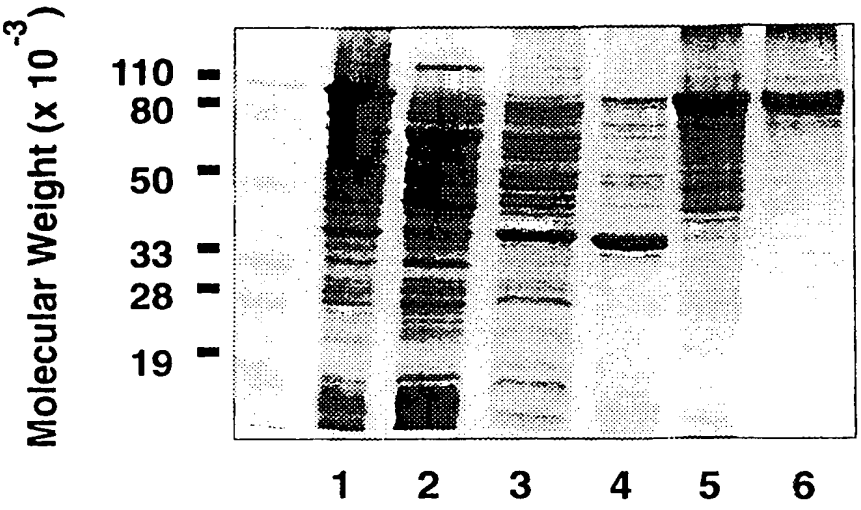
FIG.13



Purification scheme for rLbp1 expressed from E.coli

FIG.14

Purification of Q8 rLbp1 from *E. coli*



1. *E. coli* whole cells
2. Soluble proteins in 50 mM Tris/ NaCl extraction
3. Soluble proteins in Tris/ Triton X-100 extraction
4. Soluble proteins in Tris/ octylglucoside extraction
5. rLbp1 inclusion bodies
6. rLbp1

FIG.15.

FIG.16A

Nucleotide and deduced amino acid sequence of the VH19 lbpB gene.

```
MET  SER  THR  VAL  LYS  VAL  PRL  HIS  ILE  PHE
ATGAGTACTGTCAAAGTCCCCCACATTTTC
          10        20        30

TYR  GLN  LYS  ARG  THR  LEU  SER  LEU  ALA  ILE
TACCAAAAACGCACCCTTAGCCTTGCCATC
     40        50        60

ALA  SER  ILE  PHE  ALA  ALA  VAL  VAL  MET  THR
GCCAGTATTTTTGCTGCCGTGGTGATGACA
          70        80        90

GLY  CYS  ARG  SER  ASP  ASP  ILE  SER  VAL  ASN
GGCTGCCGCTCTGATGACATCAGCGTCAAT
     100       110       120

ALA  PRO  ASN  VAL  THR  GLN  LEU  PRO  GLN  GLY
GCACCCAATGTTACCCAACTGCCCCAAGGC
          130       140       150

THR  VAL  SER  PRO  ILE  PRO  ASN  THR  GLY  HIS
ACGGTTTCACCAATACCGAACACAGGTCAT
     160       170       180
```

FIG.16B

```
 ASP   ASN   THR   ASN   ASN   THR   ASN   ASN   GLN   GLY
G A C A A C A C C A A T A A C A C C A A C A A T C A G G G C
               190                 200                 210
                                    ASN   ASN   THR   ASP   ASN   SER   THR   SER   THR   THR
                                   A A C A A C A C G G A T A A C A G C A C C A G C A C A A C T
                                                       220                 230                 240

 ASP   PRO   ASN   GLY   ASP   ASN   ASN   GLN   LEU   THR
G A C C C A A A T G G C G A T A A C A A C C A A C T G A C A
               250                 260                 270
                                    GLN   ALA   GLN   LYS   THR   ALA   ALA   ALA   ALA   GLY
                                   C A A G C A C A A A A A A C T G C C G C C G C C G C A G G G
                                                       280                 290                 300

 PHE   PHE   VAL   MET   GLY   LYS   ILE   ARG   ASP   THR
T T T T T T G T G A T G G G T A A A A T T C G T G A T A C C
               310                 320                 330
                                    SER   GLU   LYS   ASN   ASP   PRO   ASP   TYR   THR   LYS
                                   A G C G A A A A A A A T G A C C C A G A T T A T A C C A A A
                                                       340                 350                 360
```

FIG.16C

```
 ASP   LEU   GLN   GLY   SER   VAL   HIS   THR   ALA   GLY
G A T T T A C A A G G C A G C G T A C A T A C A G C A G G G
              370                 380                 390
                                   GLN   GLY   LEU   GLN   TYR   LEU   GLY   THR   LYS   GLU
                                  C A A G G C T T A C A G T A C T T A G G C A C C A A A G A G
                                                 400                 410                 420

 PRO   ARG   PRO   ASP   GLY   THR   GLY   THR   GLY   LYS
C C T C G G C C A G A T G G C A C A G G T A C A G G T A A A
              430                 440                 450
                                   ASN   LEU   ARG   GLN   PRO   ILE   THR   ALA   ASP   ASP
                                  A A C T T A C G C C A G C C C A T C A C A G C T G A T G A C
                                                 460                 470                 480

 ILE   THR   PRO   LEU   TYR   PHE   ASP   LYS   PHE   PRO
A T T A C A C C A C T T T A T T T T G A T A A A T T C C C C
              490                 500                 510
                                   LYS   ILE   SER   ASP   LEU   HIS   LEU   GLU   ALS   SER
                                  A A A A T A T C C G A T C T G C A C C T A G A A A A C A G C
                                                 520                 530                 540
```

FIG.16D

```
 GLU  HIS  VAL  PHE  ASP  ALA  LYS  LYS  ALA  ASN
GAGCATGTGTTTGATGCTAAAAAAGCAAAT
          550            560            570
                          ASN  ILE  LYS  ILE  TYR  GLY  TYR  SLY  ALA  LEU
                         AACATCAAAATATATGGTTATGGTGCATTG
                                    580            590            600

 SER  SER  PRO  ALA  LYS  ASN  PRO  THR  TYR  MET
TCATCACCTGCCAAAAACCCAACCTACATG
          610            620            630
                          ASN  TYR  GLN  GLN  GLU  GLN  ASN  ILE  LYS  ASN
                         AATTATCAACAAGAACAAAAACATCAAAAAC
                                    640            650            660

 LYS  LYS  PRO  GLY  ASP  ASP  TYR  GLN  ASN  ILE
AAAAAACCAGGCGATGATTATCAAAACATT
          670            680            690
                          ARG  PHE  GLY  TYR  MET  GLU  LEU  ARG  GLU  LEU
                         CGTTTTGGCTATATGGAGCTAAGAGAGCTG
                                    700            710            720
```

FIG.16E

```
ASP  LEU  ASN  LYS  LYS  GLY  ALA  ASP  THR  GLN
G A C C T A A A T A A A A A A G G T G C A G A C A C C C C A G
                730                 740                 750

SER  ASP  LYS  ASN  ARG  ALA  ILE  ILE  PHE  THR
A G C G G A C A A G A A C C G T G C C C A T C A T T T T C A C C
                760                 770                 780

THR  PRO  THR  LEU  PHE  TYR  HIS  GLY  GLN  ASN
A C A C C C T A C C T T T T A T T T T A T C A T G G T G A G A A T
                790                 800                 810

ALA  SER  THR  HIS  LEU  PRO  LYS  ALA  GLY  LYS
G C C C A G C A C C C C A T C T G C C C A A A G G C G G G T A A A
                820                 830                 840

PHE  THR  ASP  LYS  VAL  GLY  THR  TYR  PHE
T T T G A C T A T G A G G G C A A A T T G G T T G T A T C T T G
                850                 860                 870

ASN  SER  THR  ARG  LYS  SER  ASN  GLU  GLY  ASP
A C C G A T G T C A A A A A A A C G C C C C A T T T T T A G A T
                880                 890                 900
```

FIG.16F

```
 LYS  THR  ASP  ASP  LYS  VAL  GLY  THR  TYR  PHE
A A A A C A G A C G A T A A A G T A G G C A C T T A T T T T
              910                  920                  930
                                    ASN  SER  THR  ARG  LYS  SER  ASN  GLU  GLY  ASP
                                   A A C T C A A C C A G A A A A T C A A A T G A A G G C G A T
                                                  940                  950                  960

 LEU  VAL  SER  ALA  ALA  HIS  ILE  TYR  LEU  ASN
T T G G T G A G T G C A G C A C A C A T T T A T C T A A A C
              970                  980                  990
                                    SER  PHE  LYS  TYR  LYS  HIS  THR  PRO  ALA  THR
                                   A G C T T T A A A T A T A A A C A C A C C C C G G C C A C T
                                                  1000                 1010                 1020

 TYR  SER  VAL  ASP  PHE  ASP  GLN  ASN  THR  LEU
T A T A G C G T G G A C T T T G A T C A A A A T A C C C T A
              1030                 1040                 1050
                                    LYS  GLY  LYS  LEU  SER  TYR  TYR  ASP  ASN  PRO
                                   A A A G G C A A A T T G T C T T A T T A T G A C A A C C C A
                                                  1060                 1070                 1080
```

FIG.16G

```
 ASN  LYS  GLN  THR  ALA  ASP  GLY  ARG  TYR  ILE
A A C A A G C A A A C A G C C G A T G G G C G T T A T A T C
            1090              1100              1110
                                ARG  SER  GLN  PHE  ASP  THR  ASP  LYS  LYS  VAL
                               A G A A G T C A G T T T G A T A C C G A C A A A A A G G T C
                                          1120              1130              1140

 ASN  GLU  ALA  ASP  VAL  TYR  GLU  ILE  ASP  ALA
A A T G A A G C C G A T G T C T A T G A G A T T G A C G C C
            1150              1160              1170
                                LYS  ILE  ASN  GLY  ASN  ARG  PHE  THR  GLY  THR
                               A A G A T T A A T G G C A A C C G C T T T A C T G G C A C A
                                          1180              1190              1200

 ALA  LYS  SER  LEU  ILE  ASP  ASP  ASN  THR  ASN
G C C A A A T C T T T G A T T G A T G A T A A C A C C A A T
            1210              1220              1230
                                THR  ALA  PRO  PHE  VAL  LYS  GLU  LEU  PHE  SER
                               A C C G C A C C T T T T G T T A A A G A G C T G T T C T C C
                                          1240              1250              1260
```

FIG.16H

```
 LYS  LYS  ALA  ASN  PRO  ASN  ASN  PRO  ASP  PRO
AAAAAAGCCAATCCCAACAACCCAGACCCC
          1270           1280           1290
                          ASN  SER  ASP  THR  LEU  GLU  GLY  GLY  PHE  TYR
                         AACTCAGATACGCTAGAAGGCGGGTTTTAT
                                    1300           1310           1320

 GLY  GLU  SER  GLY  ASP  GLU  LEU  ALA  GLY  LYS
GGTGAGTCGGGCGATGAGCTGGCGGGTAAA
          1330           1340           1350
                          PHE  LEU  SER  ASN  ASP  ASN  ALA  THR  PHE  VAL
                         TTTTTATCCAATGACAACGCAACTTTTGTG
                                    1360           1370           1380

 VAL  PHE  GLY  GLY  LYS  ARG  ASP  LYS  THR  THR
GTCTTTGGTGGCAAACGAGACAAAACGACC
          1390           1400           1410
                          GLU  PRO  VAL  ALA  THR  LYS  THR  VAL  THR  PHE
                         GAACCTGTCGCCACAAAAACGGTGTATTTT
                                    1420           1430           1440
```

FIG.16I

```
SER  THR  GLY  PHE  GLU  LYS  POE  SER  THR  SER
A G T A C A G G A T T T G A A A A A C C C A G C A C C C A G C
          1450                1460                1470

                PHE  VAL  GLY  ASN  GLU  GLU  ILE  GLY  SER  ILE
T T T G T T G G G C A A T G A A G A G A T T G G T A G C A T T
          1480                1490                1500

ILE  ASP  GLY  LYS  LYS  LEU  ASN  ASP  GLU  VAL
A T T G A C G G G T A A A A A G T T A A A T G A T G A A G T C
          1510                1520                1530

                ASN  ASN  GLN  ILE  GLU  ASP  GLU  THR  VAL  PRO
A A T A A T C A A A T T G A A G A T G A A A C T G T C C C T
          1540                1550                1560

VAL  SER  ASN  LYS  GLU  TYR  TYR  GLU  TYR  ASN
G T C A G T A A T A A A G A A T A T T A T G A A T A T A A T
          1570                1580                1590

                TYR  GLY  ARG  PRO  ASN  LYS  GLN  PHE  THR  LYS
T A T G G A C G A C C C C A A C A A A C A A T T C A C C A A A
          1600                1610                1620
```

FIG.16J

```
LYS  ILE  ASN  ALA  SER  VAL  GLN  LYS  ASN  PRO
A A A A T A A A C G C C A G C G T C C A A A A A A C C C C T
                1630                 1640                 1650

ALA  TYR  PHE  GLY  GLN  HIS  ASP  LYS  PHE  TYR
G C T T A T T T T G G T C A G C A T G A T A A G T T T T A T
                1660                 1670                 1680

PHE  ASN  GLY  ASN  TYR  TYR  ASP  LEU  SER  ALA
T T T A A T G G T A A C T A T T A T G A C T T A T C A G C C C
                1690                 1700                 1710

LYS  GLU  ALA  ASN  LYS  LEU  GLY  VAL  THR  ASP
A A A G A A G C A A A C A A G C T T G G T G T C T C C C A A
                1720                 1730                 1740

ASP  THR  SER  THR  ASN  LYS  SER  ILE  LEU  ALA
G A T A C C A G C A C C C A A T A A G A G T A T T T T G G C T
                1750                 1760                 1770

LYS  TYR  PRO  ASP  ALA  LYS  VAL  SER  THR  ASP
A A A T A C C C A G A T G C C A A A A G T A A G C A C A G A C
                1780                 1790                 1800
```

FIG.16K

```
 ASN  LYS  VAL  THR  LYS  ILE  VAL  LEU  GLN  GLN
AATAAAGTTACCAAAATCGTTCTACAACAA
          1810           1820           1830
                                ALA  LYS  ASP  LYS  PRO  TYR  THR  ALA  ILE  HIS
                               GCCAAAGATAAGCCGTATACCGCCATTCAT
                                          1840           1850           1860

 ALA  LYS  SER  TYR  ASP  HIS  ILE  SER  PHE  GLY
GCCAAAAGCTATGACCACATCAGTTTTGGT
          1870           1880           1890
                                GLU  VAL  LEU  TYR  ASN  ASP  ASN  LYS  GLY  ASN
                               GAAGTATTGTATAATGATAACAAAGGCAAC
                                          1900           1910           1920

 PRO  THR  ARG  SER  TYR  PHE  VAL  GLN  GLY  GLY
CCAACACGCAGTTATTTTGTGCAAGGCGGT
          1930           1940           1950
                                GLN  ALA  ASP  VAL  SER  THR  GLN  LEU  PRO  SER
                               CAAGCGGATGTCAGTACTCAGCTGCCCAGT
                                          1960           1970           1980
```

FIG.16L

```
 ALA  GLY  LYS  PHE  THR  TYR  ASN  GLY  LEU  TRP
G C A G G T A A A T T C A C C T A T A A T G G T C T T T G G
              1990              2000              2010
                                 ALA  GLY  TYR  LEU  THR  GLN  LYS  LYS  ASP  LYS
                                G C A G G C T A C C T G A C C C A G A A A A A A G A C A A A
                                                  2020              2030              2040

 GLY  TYR  SER  LYS  ASP  GLU  ASP  THR  ILE  LYS
G G T T A T A G C A A A G A T G A G G A T A C C A T C A A G
              2050              2060              2070
                                 GLN  LYS  GLY  LEU  LYS  ASP  TYR  ILE  LEU  THR
                                C A A A A A G G T C T T A A A G A T T A T A T A T T G A C C
                                                  2080              2090              2100

 LYS  ASP  PHE  ILE  PRO  GLN  ASP  ASP  ASP  ASP
A A A G A C T T T A T C C C A C A A G A T G A C G A T G A C
              2110              2120              2130
                                 ASP  ASP  ASP  SER  LEU  THR  ALA  SER  ASP  ASP
                                G A T G A C G A T A G T T T G A C C G C A T C T G A T G A T
                                                  2140              2150              2160
```

FIG.16M

```
 SER  GLN  ASP  ASP  ASN  THR  HIS  GLY  ASP  ASP
TCACAAGATGATAATACACATGGCGATGAT
          2170           2180           2190
                          ASP  LEU  ILE  ALA  SER  ASP  ASP  SER  GLN  ASP
                         GATTTGATTGCATCTGATGATTCACAAGAT
                                     2200           2210           2220

 ASP  ASP  THR  ASP  GLY  ASP  ASP  ASP  SER  ASP
GATGACACAGATGGCGATGACGATTCAGAT
          2230           2240           2250
                          ASP  LEU  GLY  ASP  GLY  ALA  ASP  ASP  ASP  ALS
                         GATTTGGGTGATGGTGCAGATGATGACGCC
                                     2260           2270           2280

 ALA  GLY  LYS  VAL  TYR  HIS  ALA  GLY  ASN  ILE
GCAGGCAAAGTGTATCATGCAGGTAATATT
          2290           2300           2310
                          ARG  PRO  GLU  PHE  GLU  ASN  LYS  TYR  LEU  PRO
                         CGCCCTGAATTTGAAAACAAATACTTGCCC
                                     2320           2330           2340
```

FIG.16N

```
 ILE   ASN   GLU   PRO   THR   HIS   GLU   LYS   THR   PHE
A T T A A T G A G C C T A C T C A T G A A A A A A C C T T T
            2350                2360                2370
                                  ALA   LEU   ASP   GLY   LYS   ASN   LYS   ALA   LYS   PHE
                                  G C C C T A G A T G G T A A A A A T A A G G C T A A G T T T
                                              2380                2390                2400

 ASP   VAL   ASN   PHE   ASP   THR   ASN   SER   LEU   THR
G A T G T A A A C T T T G A C A C C A A C A G C C T A A C T
            2410                2420                2430
                                  GLY   LYS   LEU   ASN   ASP   GLU   ARG   GLY   ASP   ILE
                                  G G T A A A T T A A A C G A T G A G A G A G G T G A T A T C
                                              2440                2450                2460

 VAL   PHE   ASP   ILE   LYS   ASN   GLY   LYS   ILE   ASP
G T C T T T G A T A T C A A A A A T G G C A A A A T T G A T
            2470                2480                2490
                                  GLY   THR   GLY   PHE   THR   ALA   LYS   ALA   ASP   VAL
                                  G G C A C A G G A T T T A C C G C C A A A G C C G A T G T G
                                              2500                2510                2520
```

FIG.16O

```
 PRO  ASN  TYR  ARG  GLU  GLU  VAL  GLY  ASN  ASN
C C A A A C T A T C G T G A A G A A G T G G G T A A C A A C
            2530              2540              2550
                                 GLN  GLY  GLY  GLY  PHE  LEU  TYR  ASN  ILE  LYS
                                C A A G G T G G C G G T T C T T A T A C A A C A T C A A A
                                            2560              2570              2580

 ASP  ILE  ASP  VAL  LYS  GLY  GLN  PHE  PHE  GLY
G A T A T T G A T G T T A A G G G G C A A T T T T T T G G C
            2590              2600              2610
                                 THR  ASN  GLY  GLU  GLU  LEU  ALA  GLY  ARG  LEU
                                A C A A A T G G C G A A G A G T T G G C A G G A C G G T T A
                                            2620              2630              2640

 HIS  HIS  ASP  LYS  GLY  ASP  GLY  ILE  THR  ASP
C A T C A T G A C A A A G G C G A T G G C A T C A C T G A C
            2650              2660              2670
                                 THR  ALA  GLU  LYS  ALA  GLY  ALA  VAL  PHE  GLY
                                A C C G C C G A A A A A G C A G G G G C T G T C T T T G G G
                                            2680              2690              2700

 ALA  VAL  LYS  ASP  LYS  ***
G C T G T T A A A G A T A A A T A A
            2710
```

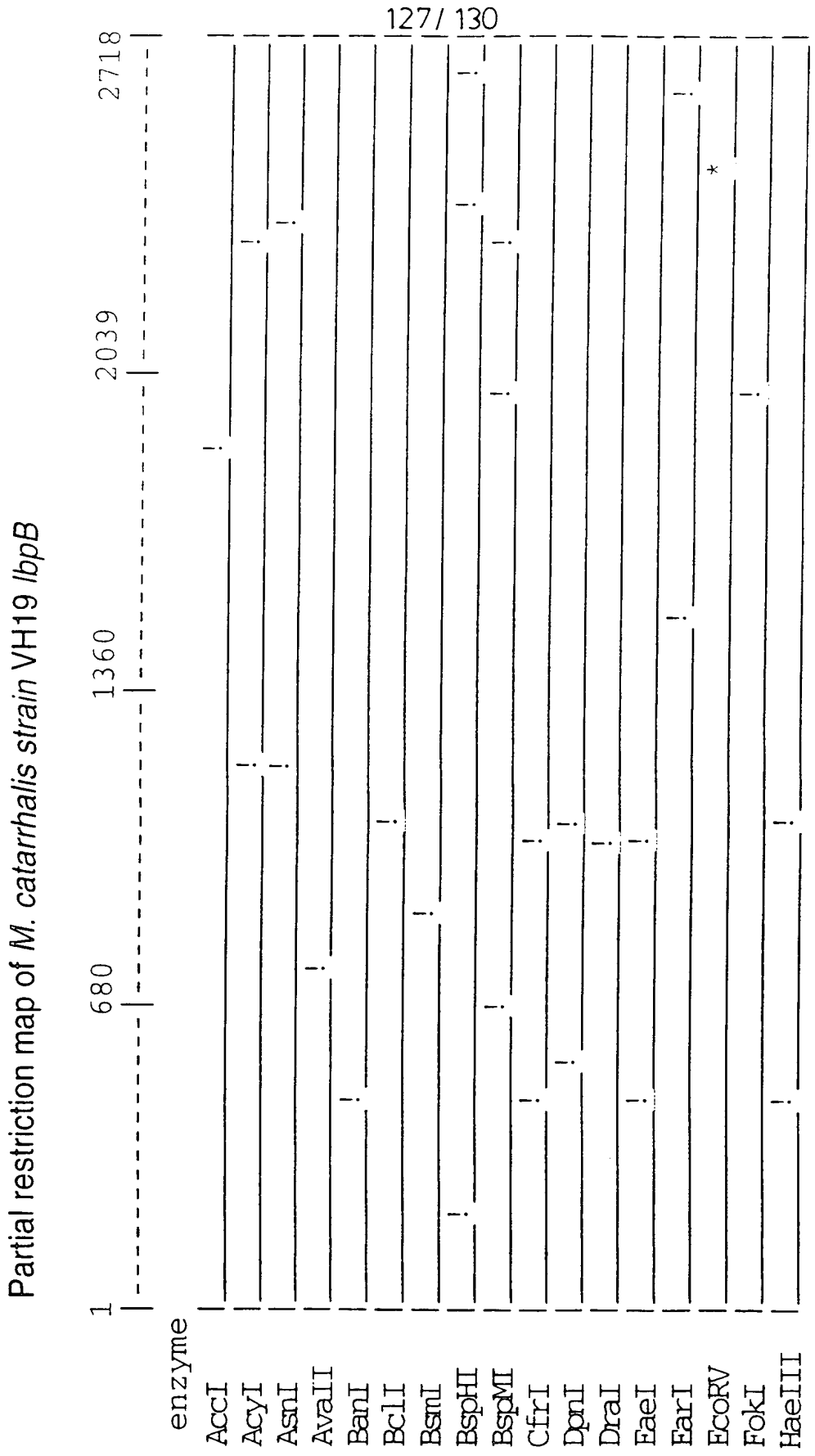


FIG.17A

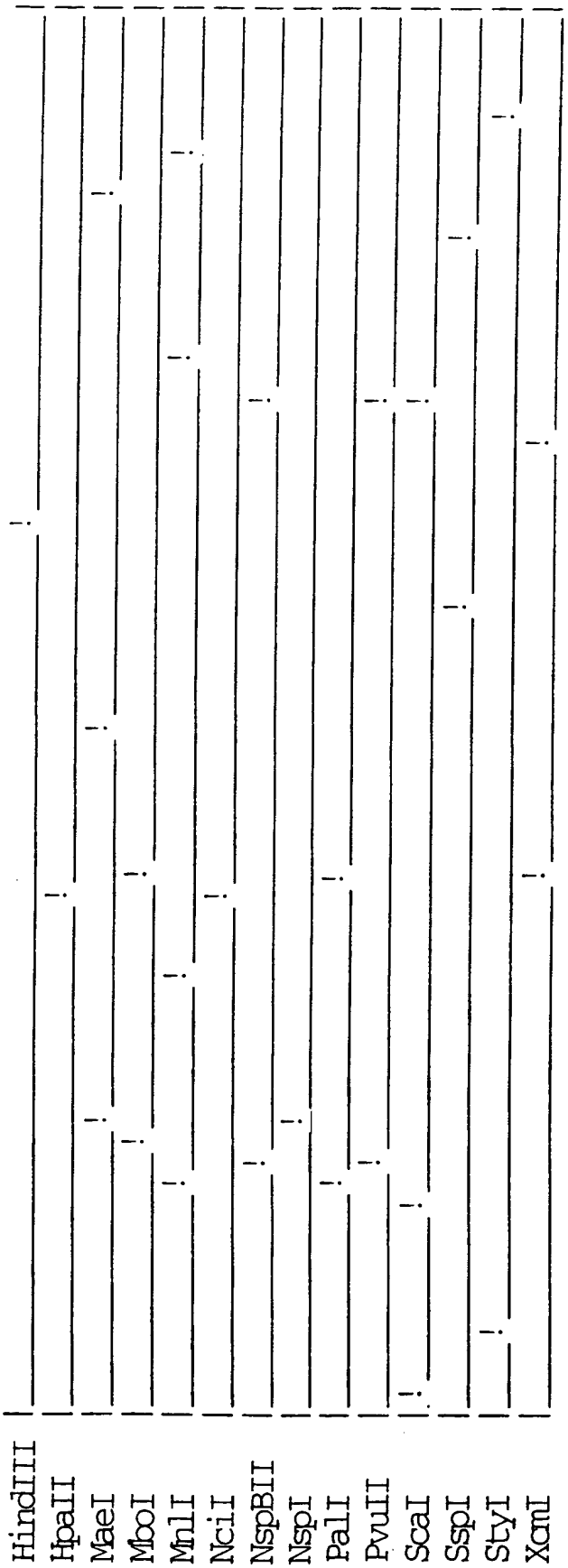

HindIII
HpaII
MaeI
MboI
MnlI
NciI
NspBII
NspI
PalI
PvuII
ScaI
SspI
StyI
XcmI

FIG.17B

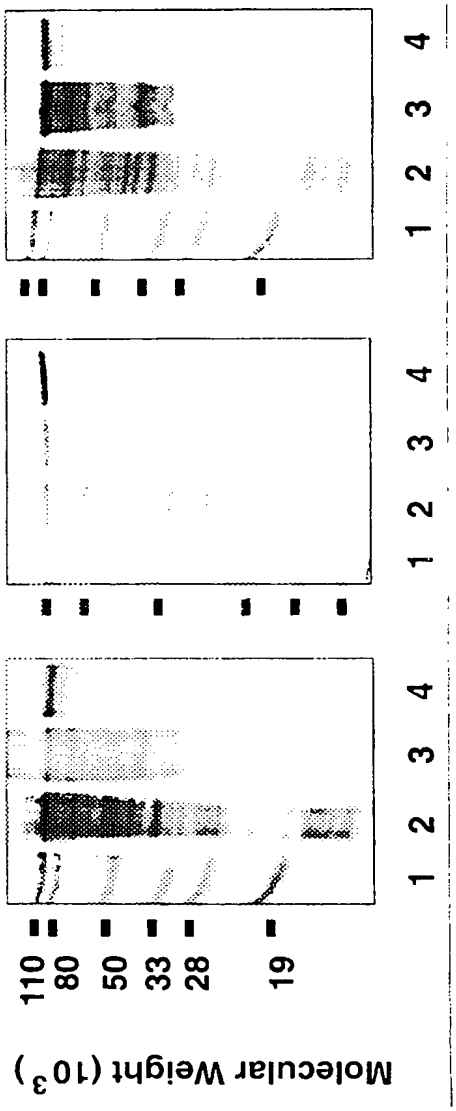


FIG.18A.

FIG.18B.

FIG.18C.

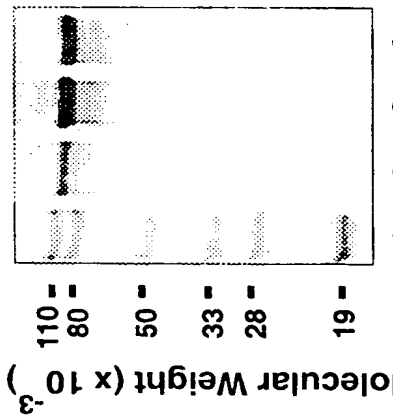


FIG.19A.

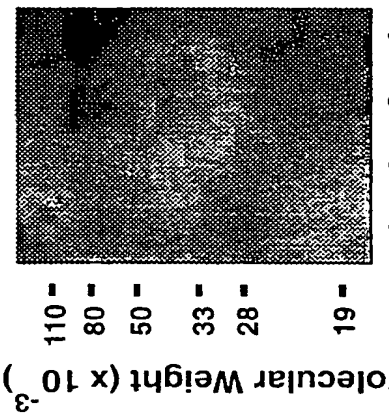


FIG.19B.

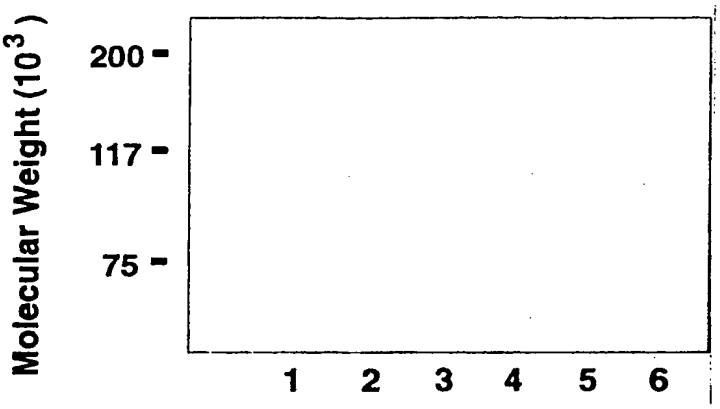


FIG.20A.

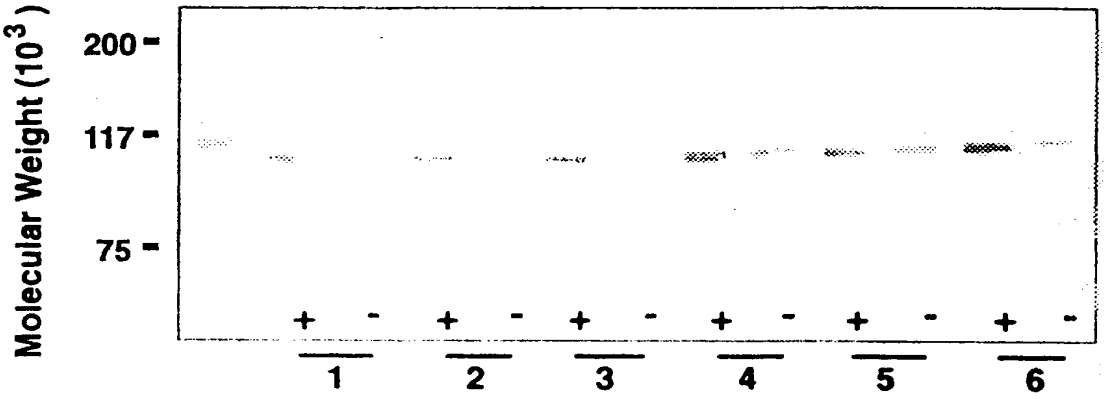


FIG.20B.

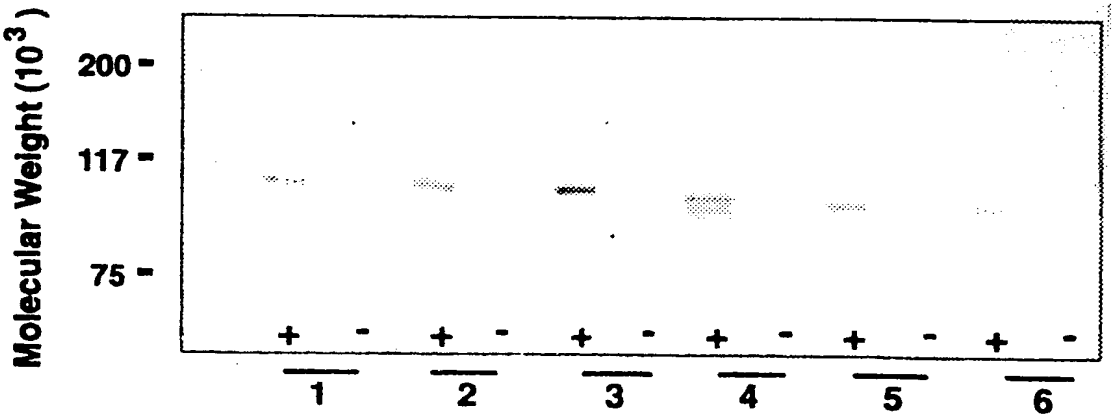


FIG.20C.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/79, 14/22,
A61K 38/17, G01N 33/68, C12Q 1/68

A3

(11) International Publication Number: WO 98/55606

(43) International Publication Date: 10 December 1998 (10.12.98)

(21) International Application Number: PCT/CA98/00544

(22) International Filing Date: 2 June 1998 (02.06.98)

(30) Priority Data:

08/867,941	3 June 1997 (03.06.97)	US
09/074,658	8 May 1998 (08.05.98)	US

(71) Applicant *(for all designated States except US)*: CONNAUGHT LABORATORIES LIMITED [CA/CA]; 1755 Steeles Avenue West, North York, Ontario M2R 3T4 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LOOSMORE, Sheena, M. [CA/CA]; 70 Crawford Rose Drive, Aurora, Ontario L4G 4R4 (CA). DU, Run-Pan [CA/CA]; (CA). WANG, Quijun [CA/CA]; 299 Chelwood Drive, Thornhill, Ontario L4J 7Y8 (CA). YANG, Yan-Ping [CA/CA]; Apartment 709, 120 Torresdale Avenue, Willowdale, Ontario M2R 3N7 (CA). KLEIN, Michel, H. [CA/CA]; 16 Munro Boulevard, Willowdale, Ontario M2P 1B9 (CA).

(74) Agent: STEWART, Michael, I.; Sim & McBurney, 6th floor, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
4 March 1999 (04.03.99)

(54) Title: LACTOFERRIN RECEPTOR GENES OF MORAXELLA

(57) Abstract

Purified and isolated nucleic acid molecules are provided which encode lactoferrin receptor proteins of *Moraxella*, such as *M. catarrhalis*, or a fragment or an analog of the lactoferrin receptor protein. The nucleic acid sequence may be used to produce recombinant lactoferrin receptor proteins Lbp1, Lbp2 or ORF3 of the strain of *Moraxella* free of other proteins of the *Moraxella* strain for purposes of diagnostics and medical treatment. Furthermore, the nucleic acid molecule may be used in the diagnosis of infection.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Interi nal Application No
PCT/CA 98/00544

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/79 C07K14/22 A61K38/17 G01N33/68 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N A61K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CA 2 162 193 A (CONNAUGHT LAB) 3 May 1997	20-22, 24,29,30
Y	see the claims see abstract; figures 1-3; examples 1,4 see page 4, line 21 - page 7, line 15	3,4, 6-19,23, 25-28
X	OGUNNARIWO J A AND SCHRYVERS A B: "Rapid identification of bacterial transferrin and lactoferrin receptor protein genes" JOURNAL OF BACTERIOLOGY, vol. 178, no. 24, December 1996, pages 7326-7328, XP002083300	1,2,5, 31,32
Y	see abstract; figure 1; table 1 see page 7327 - page 7328	3,4, 6-19,23, 25-28

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the pnority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 November 1998	23/12/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Oderwald, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interr ...al Application No
PCT/CA 98/00544

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	WO 90 12591 A (UNIV TECHNOLOGIES INT ;SCHRYVERS ANTHONY BERNARD (CA)) 1 November 1990 see abstract; claims 5-8,16-26; figure 1; examples 1,2 see page 5, line 4 - page 6, line 26 see page 7, line 30 - page 8, line 1 see page 10, line 8 - page 11, line 8 ---	20,24,29
X	SCHRYVERS A B ET AL: "COMPARATIVE ANALYSIS OF THE TRANSFERRIN AND LACTOFERRIN BINDING PROTEINS IN THE FAMILY NEISSERIACEAE" CANADIAN JOURNAL OF MICROBIOLOGY, vol. 35, no. 5, May 1989, pages 409-415, XP002020995 cited in the application see abstract; figures 1,4-6 see page 410, paragraph 9 see page 414, paragraph 2 ---	20-22,24
T	DU R-P ET AL: "Cloning and expression of the Moraxella catarrhalis lactoferrin receptor genes" INFECTION AND IMMUNITY, vol. 66, no. 8, August 1998, pages 3656-3665, XP002083301 see the whole document -----	1-32

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA 98/ 00544

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim(s) 30
is(are) directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please see additional sheet.

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [X] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1, 5-20, 24, 29-32 all partially and 2, 21.

   Lbp1 polypeptide, corresponding nucleic acids, vectors containing said nucleic acid, host cells transformed with said vector, immunogemmic composition containing said nucleic acid or protein, methods for detection and diagnostic kit involving said nucleic acid.

2. Claims: 1, 5-20, 24, 29-32 all partially and 3, 22.

   Same as in invention 1 but comprising Lbp2.

3. Claims: 1, 5-20, 24, 29-32 all partially and 4, 23, 25-28.

   Same as in invention 1 but comprising ORF3.

INTERNATIONAL SEARCH REPORT

...formation on patent family members

Inter nal Application No
PCT/CA 98/00544

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CA 2162193 A	03-05-1997	NONE	
WO 9012591 A	01-11-1990	US 5292869 A	08-03-1994
		AU 649950 B	09-06-1994
		AU 5526190 A	16-11-1990
		EP 0528787 A	03-03-1993
		JP 4506794 T	26-11-1992
		NZ 247967 A	24-02-1995
		US 5141743 A	25-08-1992